SCIT

School of Computing and Information Technology
Faculty of Engineering & Information Sciences

SIM Session 1, 2022
Subject Outline
INFO411 – Data Mining and Knowledge Discovery

Subject Organisation

Subject Coordinator/Lecturer:	Assoc. Prof. Wanqing Li
Email:	wanqing@uow.edu.au
Credit Points:	6 credit points
Duration:	1 session
Lecture Times & Location:	Refer to SIMConnect

The University uses the eLearning system Moodle to support all coursework subjects.

Students should check the subject's Moodle site regularly as important information, including **details of unavoidable changes in assessment requirements will be posted from time to time** http://www.uow.edu.au/student/. Any information posted to the web site is deemed to have been notified to all students.

In extraordinary circumstances the provisions stipulated in this Subject Outline may require amendment after the Subject Outline has been distributed. All students enrolled in the subject must be notified and have the opportunity to provide feedback in relation to the proposed amendment, prior to the amendment being finalised.

Data on student performance and engagement (such as Moodle and University Library usage, task marks, use of SOLS) will be available to the Subject Coordinator to assist in analysing student engagement, and to identify and recommend support to students who may be at risk of failure. If you have questions about the kinds of data the University uses, how we collect it, and how we protect your privacy in the use of this data, please refer to http://www.uow.edu.au/dvca/bala/analytics/index.html

Subject Description

Introduction to Data Mining, Knowledge Discovery, and Big Data with coverage of Data Structures, role of Data Quality and per-processing, Association Rules, Artificial Neural Networks, Support Vector methods, Tree Based Methods, Clustering and Classification Methods, Regression and Statistical Methods, Overfitting and Inferential issues, Evaluation, Use of Data Mining packages with applications for benchmark and real world situations.

Subject Learning Outcomes

On successful completion of this subject, students will be able to:

1. Identify useful relationships and important subgroups in large data sets;
2. Suggest appropriate approaches and solutions to given data mining problems;
3. Plan and carry out analyses of large and complex data sets;
4. Use parametric, non-parametric, and probabilistic methods to model data in various domains;
5. Analyse and interpret results;
6. Use data mining software such as R as well as use relevant plugins and software packages;
7. Analyse data mining algorithms and techniques;
8. Understand the role and challenges of methods in Big Data applications;
9. Identify and distinguish data mining applications from other IT applications;
10. Describe data mining algorithms;
11. Compare the applicability of data mining applications.

Recent Improvements

The School is committed to continual improvement in teaching and learning and takes into consideration student feedback from many sources. These sources include direct student feedback to tutors and lecturers, feedback through Student Services and the Faculty Central, and responses to the Subject Evaluation Surveys. This information is also used to inform comprehensive reviews of subjects and courses.

Attendance Requirements

It is the responsibility of students to attend all lectures/tutorials/labs/seminars/practical work for subjects for which you are enrolled.

Satisfactory attendance is deemed by the University, to be attendance at approximately 80% of the allocated contact hours.

Method of Presentation

The subject will be presented as a series of lectures and laboratories.

Students must be aware that they are responsible for their own learning. Students must prepare adequately for lectures and tutorials in order to properly digest the material presented in those forms. Students are expected to undertake private study in order to fully understand and integrate all the material covered in this unit.

Subject Material

Any readings/references are recommended only and are not intended to be an exhaustive list. Students are encouraged to use the library catalogue and databases to locate additional readings.

Recommended Textbook

Introduction to Data Mining (2nd Edition), Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar, Pearson, 2018, ISBN-13: 978-0133128901.

Reference Books

1. A. B. M. Shawkat Ali, Saleh A. Wasimi, "Data Mining:Methods and Techniques", Thomson, 2007, ISBN 978-0-17-013676-1
2. Ian H. Witten, Eibe Frank "Data Mining Practical Machine Learning Tools and Techniques", Elsevier inc., 2005, ISBN 0-12-088407-0
3. Jiawei Han, Micheline Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann publishers, 2006, ISBN 978-1-55860-901-3
4. Margaret Dunham, "Data Mining Introductory and Advanced Topics, Pearson Education Inc., 2003, ISBN0-13-088892-3
5. Graham Williams, "Data Mining with Rattle and R: the art of excavating data for knowledge discovery", Springer Verlag, 2011, ISBN 9781441998903

Assessment

This subject has the following assessment components.

ASSESSMENT ITEMS & FORMAT	% OF FINAL MARK	GROUP/ INDIVIDUAL	DUE DATE	SUBJECT LEARNING OUTCOMES	CRITERIA TO ASSESS ITEM
Assignment 1	12%	Individual	TBD	1,3,4,5,8,11	Correctness, completeness, and consistency with specification
Assignment 2	12%	Individual	TBD	1,3,4,5,8,11	Correctness, completeness, and consistency with specification
Project	16%	Group	TBD	1,2,3,4,5,7,8, 10,11	Correctness, completeness, and

					consistency with specification
Final Examination	60%	Individual	During Exam Period	8,9,10,11	Correctness, completeness, and consistency with specification

Notes on Assessment

Plagiarism of any part of an assessed item (a lab, assignment, or project) will result in zero marks being recorded for the affected assessment.

Method of Submission of Assessment Items

All assessments are to be submitted via Moodle. Submission via email is not acceptable.

Arrangement for acknowledging submission of written work

Acknowledgement of submission will occur electronically through Moodle.

Procedure for the return of assessment items

Assessment of assignments will be done electronically on Moodle.

Procedure for the retention of assessed work

The University may retain copies of student work in order to facilitate quality assurance of assessment processes, in support of the continuous improvement of assessment design, assessment marking and for the review of the subject. The University retains records of students' academic work in accordance with the University Records Management Policy and the State Records Act 1988 and uses these records in accordance with the University Privacy Policy and the Privacy and Personal Information Protection Act 1998.

Assessment General

Submission of assessment items via email will not be accepted.

Student contributions to tutorial and/or seminars

Not applicable.

Assessment task is set up to be checked by Turnitin

This subject does not use Turnitin.

Assessment Quality Cycle

The University of Wollongong is committed to the quality assurance and quality enhancement of assessment. The University will meet its legislative and regulatory obligations, to ensure consistent and appropriate assessment through course management and coordination, including assessment quality

assurance procedures. An Assessment Quality Cycle is used to describe quality assurance at the points of assessment design, assessment delivery, the declaration of marks and grades, and review and improvement activities.

Referencing System

The type of referencing system to be used for written work is as follows:

- the Author-Date (Harvard) referencing system is the University's default referencing system to be used in the absence of a documented faculty/school preferred referencing style. Refer to the Library Referencing and Citing link: https://www.uow.edu.au/student/learningcoop/referencingciting/index.html

Internet Resources

Not applicable.

Technical Fail

To be eligible for a Pass in this subject a student must achieve a mark of at least 40% in the final exam. Students who fail to achieve this minimum mark & would have otherwise passed may be given a TF (Technical Fail) for this subject.

Penalties for late submission of assessment items

Penalties apply to all late work, except if student academic consideration has been granted. Late submissions will attract a penalty of 25% of the assessment mark per day.

This amount is per day including weekends.

Work more than 4 days late will be awarded a mark of zero.

UOW Grade Descriptors

GRADE	DESCRIPTOR
High Distinction(HD) 85-100%	For performance that provides evidence of an outstanding level of attainment of the relevant subject learning outcomes, demonstrating the attributes of a distinction grade plus (as applicable) one or more of the following: • consistent evidence of deep and critical understanding • substantial originality and insight in identifying, generating and communicating competing arguments, perspectives or problem-solving approaches • critical evaluation of problems, their solutions and their implications • use of quantitative analysis of data as the basis for deep and thoughtful judgments, drawing insightful, carefully qualified conclusions from this work • creativity in application as appropriate to the discipline

	• eloquent and sophisticated communication of information and ideas in terms of the conventions of the discipline • consistent application of appropriate skills, techniques and methods with outstanding levels of precision and accuracy • all or almost all answers correct, very few or none incorrect
Distinction (D) 75-84%	For performance that provides evidence of a superior level of attainment of the relevant subject learning outcomes, demonstrating the attributes of a credit grade plus (as applicable) one or more of the following: • evidence of integration and evaluation of critical ideas, principles, concepts and/or theories • distinctive insight and ability in applying relevant skills, techniques, methods and/or concepts • demonstration of frequent originality in defining and analysing issues or problems and providing solutions • fluent and thorough communication of information and ideas in terms of the conventions of the discipline • frequent application of appropriate skills, techniques and methods with superior levels of precision and accuracy • most answers correct, few incorrect
Credit (C) 65-74%	For performance that provides evidence of a high level of attainment of the relevant subject learning outcomes, demonstrating the attributes of a pass grade plus (as applicable) one or more of the following: • evidence of learning that goes beyond replication of content knowledge or skills • demonstration of solid understanding of fundamental concepts in the field of study • demonstration of the ability to apply these concepts in a variety of contexts • use of convincing arguments with appropriate coherent and logical reasoning • clear communication of information and ideas in terms of the conventions of the discipline • regular application of appropriate skills, techniques and methods with high levels of precision and accuracy • many answers correct, some incorrect
Pass (P) 50-64%	For performance that provides evidence of a satisfactory level attainment of the relevant subject learning outcomes, demonstrating (as applicable) one or more of the following: • knowledge, understanding and application of fundamental concepts of the field of study • use of routine arguments with acceptable reasoning

	• adequate communication of information and ideas in terms of the conventions of the discipline • ability to apply appropriate skills, techniques and methods with satisfactory levels of precision and accuracy • a combination of correct and incorrect answers
Fail (F) <50%	For performance that does not provide sufficient evidence of attainment of the relevant subject learning outcomes.
Technical Fail (TF)	When minimum performance level requirements for at least one assessment item in the subject as a whole has not been met despite the student achieving at least a satisfactory level of attainment of the subject learning outcomes.

https://www.uow.edu.au/curriculum-transformation/aqc/uowgradedescriptors/index.html

Plagiarism - University's Academic Integrity Policy

The University's policy on acknowledgement practice and plagiarism provides detailed information about how to acknowledge the work of others: http://www.uow.edu.au/about/policy/UOW058648.html

The University's Academic Integrity Policy, Faculty Handbooks and subject guides clearly set out the University's expectation that students submit only their own original work for assessment and avoid plagiarising the work of others or cheating. Re-using any of your own work (either in part or in full) which you have submitted previously for assessment is not permitted without appropriate acknowledgement or without the explicit permission of the Subject Coordinator. Plagiarism can be detected and has led to students being expelled from the University.

The use by students of any website that provides access to essays or other assessment items (sometimes marketed as 'resources'), is extremely unwise. Students who provide an assessment item (or provide access to an assessment item) to others, either directly or indirectly (for example by uploading an assessment item to a website) are considered by the University to be intentionally or recklessly helping other students to cheat. Uploading an assessment task, subject outline or other course materials without express permission of the university is considered academic misconduct and students place themselves at risk of being expelled from the University.

When you submit an assessment task, you are declaring the following

1. It is your own work and you did not collaborate with or copy from others.
2. You have read and understand your responsibilities under the University of Wollongong's Academic Integrity Policy on plagiarism.
3. You have not plagiarised from published work (including the internet). Where you have used the work from others, you have referenced it in the text and provided a reference list at the end to the assignment.

Students must remember that:

- Plagiarism will not be tolerated.
- Students are responsible for submitting original work for assessment, without plagiarising or cheating, abiding by the University's Academic Integrity Policy as set out in the University

Handbook, the University's online Policy Directory and in Faculty handbooks and subject guides.

Student Academic Complaints Policy (Coursework or Higher Degree Research)

In accordance with the Coursework Student Academic Complaints Policy, a student may request an explanation of a mark for an assessment task or a final grade for a subject consistent with the student's right to appropriate and useful feedback on their performance in an assessment task. Refer to the Coursework Student Academic Complaints Policy for further information
http://www.uow.edu.au/about/policy/UOW058653.html

General Advice

This outline should be considered in conjunction with policy documents available through the University of Wollongong website. Those policies are subject to revision.

Please see the additional documentation provided with this subject outline.