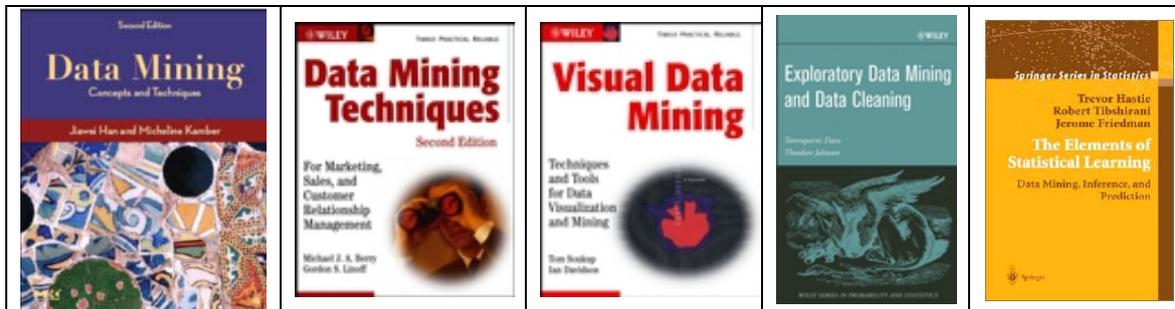# MIST812 – Decision Support Systems

Unit Outline
Second Semester 2008

Unit Convenor: Dr. Ayse Bilgin

Students in this unit should read this unit outline carefully at the start of semester. It contains important information about the unit. If anything in it is unclear, please consult the teaching staff in the unit.

ABOUT THIS UNIT

Decision support systems provide information to support semi-structured and unstructured decisions in modern organisations. The decision support systems studied in this unit are computer based and provide for user manipulation of source data extracted different sources, for example from databases both internal and external to the organisation. Students will study decision support systems involving: data warehousing and data marts, online analytic processing, data mining and the geographic information systems. The unit will be taught using modern software including SPSS Clementine for data mining and MapInfo for the GIS package.

MIST812 offered by the Statistics department in the Division of Economic and Financial Studies. This unit has been designed as a 4 postgraduate credit point unit, requiring the equivalent of 13 weeks of work over one semester. Being a four credit-points unit you should expect to spend a minimum of 12 hours per week to meet the requirements of the unit.

This unit expands on topics covered in MIST800 *Computer Applications in Business and a*vailable to graduate students in Business. It is offered as part of the Master of Commerce/Master of International Business Program.

Software:
SPSS, SPSS-CLEMENTINE, MAPINFO, Microsoft Excel
http://goto.mapinfo.com/webtutorials/
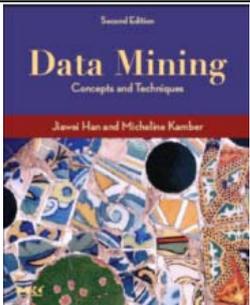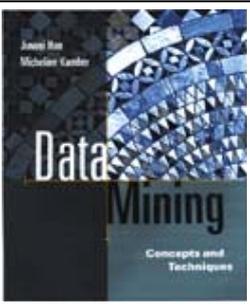
Prerequisite:
Basic quantitative knowledge (STAT170 or MIST800 or equivalent).

TEACHING STAFF

Dr Ayse Bilgin
Room: E4A 515
Phone: 9850 8509
e-mail: abilgin@efs.mq.edu.au

The best way to contact the teaching staff in this unit is through the Blackboard Mail tool.

RECOMMENDED TEXT BOOK

Data Mining: Concepts and techniques by Jiawei Han and Micheline Kamber, 2006 or 2001, Morgan and Kaufmann (library call number QA76.9.D343.H36 2001)

1) Data mining techniques for marketing, sales and customer relationship management by Michael Berry and Gordon Linoff, 2004, John Wiley (library call number HF5415.125 .B47 2004)

2) Visual Data Mining: Techniques and Tools for Data Visualization and Mining by Tom Soukup, Ian Davidson, May 2002, Wiley and Sons (library call number QA76.9.D343 S68 2002)

3) Exploratory Data Mining and Data Cleaning by Tamraparni Dasu, Theodore Johnson, May 2003 (library call number QA76.9.D343 D34 2003)

4) The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Trevor HASTIE, Robert TIBSHIRANI,and Jerome FRIEDMAN. New York: Springer-Verlag, 2001. ISBN 0-387-95284-5. (library call number Q325.75.F75 2001)

RECOMMENDED TEXT BOOKS – ONLINE

1) CRoss Industry Standard Process for Data Mining
http://www.crisp-dm.org/download.htm

2) Introduction to Data Mining and Knowledge Discovery
http://www.twocrows.com/intro-dm.pdf

WEEKLY READINGS

The weekly readings are available through Blackboard.

CLASSES

Lectures
Lectures begin in Week 1. Students should attend **ONE** 2-hour session per week: Mondays between 6:00 and 8:00pm in E7B164.

Tutorials
Tutorials also begin in Week 1. The aim of tutorials is to practise techniques learnt in lectures. They are designed so that students work through the exercises and ask as many questions as they need to improve their understanding. Tutors are the facilitators in the tutorial groups. They will assist students and instead of giving them straight answers for every question, they will create an environment for thinking process and discussion between the students. Tutorials will be held in Statistical Computing Lab: **E4B 3O8** after the lecture Mondays between 8:00 and 10:00pm.

The timetable for classes can be found on the University web site at:
http://www.timetables.mq.edu.au/

UNIT WEB PAGE

Information relating to this unit can be found by visiting the Macquarie University Statistics Department web site. The URL for this unit is

http://www.stat.mq.edu.au/pg/stat_units/mist/mist812

BLACKBOARD ACCESS

There is a Blackboard site for this subject. Students are required to log into Blackboard using their Student ID Number and myMQ Portal Password. The web site for all online units is http://learn.mq.edu.au which includes useful links regarding study tips and technical information.

If you have a personal question, please send an e-mail to the lecturer through Blackboard e-mail facility or alternatively you could send a normal e-mail by using your Macquarie University student e-mail account.

The lecturers will make announcements via the Blackboard. Accordingly, you should make sure you log in and read the posts at least twice a week.

LEARNING OUTCOMES

On completion of this unit participants will be able to:

- gained theoretical and practical experience of the principles and the concepts of the data mining
- know how to create concept hierarchies
- use market basket analysis to improve the sales of a given company
- use classification and cluster analysis as data mining tools
- understand how the decision trees are developed and be able to interpret the outputs of the decision trees
- organise data which is suitable to display as a map
- create active earth maps with hyperlinks
- use MapInfo Software to create region boundaries for a chosen region
- to understand the link between data mining, GIS and good decision making
- write (data mining) report/s for their analysis

GENERIC SKILLS

In addition to discipline specific learning outcomes, all academic programs at Macquarie assist students to develop or enhance their generic skills in a range of areas. In this unit, participants will have the opportunity to practice and/or further develop their:

- Written and oral communication skills, particularly report writing skills,
- Skills in identifying, analysing and solving problems,
- Ability to apply theory to practice,
- Improve their ability to work co-operatively as a team member,
- Enhance their critical thinking skills,
- Be confident in the use of a variety of software packages for solving problems, and
- Self and peer assessment skills.

TEACHING AND LEARNING STRATEGY

- students are expected to attend thirteen weeks of the lectures and the tutorials
- in data mining section, weekly practical exercises are set for individual (formative) assessment of their learning. However, students can work in groups, but the submitted work must be the students' own work (and words)
- Lab exercises can be submitted electronically (Blackboard) or as hard copy to lecturer.
- the GIS project will be prepared as a report and presented in week 13
- if for any reason, students can not hand in their Assignments on time, they have to contact the lecturer in advance

RELATIONSHIP BETWEEN ASSESSMENT AND LEARNING OUTCOMES

While attendance at classes is important, it is only a small proportion of the total workload for the unit: reading, research in the library, working with other students in groups, completing assignments, using the computer and private study are all components of the work involved. At Macquarie it is expected that the average student should spend three hours per week per credit point.

You are expected to present yourself for examination at the time and place designated in the University Examination Timetable. The only exception to not sitting an examination at the designated time is because of documented illness or unavoidable disruption. In these circumstances you may wish to consider applying for Special Consideration. Information about unavoidable disruption and the special consideration process is available at

http://www.reg.mq.edu.au/Forms/APSCons.pdf and
http://www.efs.mq.edu.au/student_support/important_processes

If a Supplementary Examination is granted as a result of the Special Consideration process, the examination will be scheduled after the conclusion of the official examination period.

You are advised that it is Macquarie University policy not to set early examinations for individuals or groups of students. All students are expected to ensure that they are available until the end of the teaching semester that is the final day of the official examination period.

ASSESSMENT

For each assessment task (lab exercises and/or individual assignments), your student ID, name and the assessment identifier should appear in the header and/or footer of every page of your submission. When naming your files please adopt the following convention:

**StudentID-(Your Surname)(Initial of Your First Name) − Lab Week 2
e.g., 40000000-BilginA-LabWeek 2 or
40000000-BilginA-GIS Project**

No other format of naming the assessment tasks will be accepted. If it is not possible to submit an assessment electronically, it can be submitted to LIC.

Your files should be saved as Word 2003 or pdf before uploaded in Blackboard or e-mailed to LIC (unless specified otherwise by LIC).

The cover sheets for the individual work and group work can be downloaded from http://www.efs.mq.edu.au/EFS_docs/student_support/Individual_cover_sheet.pdf and http://www.efs.mq.edu.au/EFS_docs/student_support/Group_cover_sheet.pdf respectively, however they are not needed when submitting your files electronically.

Seven weekly lab exercises related to data mining are due at BEGINNING of your lecture session on week following date of issue (e.g. Week 2 lab exercise solution is due in Week 3 before the lecture or by 6pm). You need to submit them electronically before the lecture or you can hand them to Dr Bilgin at the beginning of the lecture. These weekly lab exercises will be corrected by your tutor. They are formative assessment tools which mean that there will not be any marks for these weekly exercises that goes toward your final grade. However the feedback(s) provided will be valuable for your learning. NO LATE SUBMISSIONS WILL BE ACCEPTED.

Data mining projects will be done in a group of maximum five students. There will be two projects, the first is worth 10% and the second is worth 20% of total assessment marks. The second project is built on the results of the first project.

GIS project will be worth 10% of total assessment marks. You will need to work individually. Each student will present (5 or 10 minutes) their project and will be graded on their presentation. Presentation time table will be available from Blackboard.

Final examination (60%) is 3 hours long with 10 minutes reading time. This exam will be held during the examination period and will examine any material covered throughout the course. The examination is 'closed book'. You may refer only to a single self-prepared **hand-written** A4 sheet of crib notes which may be written on both sides. This summary must be submitted with your exam paper and is marked for conforming to the guidelines given. Any other materials such as lecture notes and textbooks are not permitted.

Calculators are permitted, but may be used only as calculators, and not as storage devices. Mobile phones should be disabled.

To satisfactorily complete the course work part of this unit you will be expected to:

- attend the lectures and tutorials,
- submit at least 70% of the lab exercises,
- submit three assignments (two DM and one GIS) on or before the due dates.

To obtain a passing grade, both coursework and exam performance must be satisfactory.

OVERALL ASSESSMENT

The assessment requirements will contribute to your overall mark and grade for the unit according to the following weightings:

Data Mining Project 1	10%
Data Mining Project 2	20%
GIS Project Report	5%
GIS Project Presentation	5%
Final Examination	60%

The mark (SNG) recorded for this unit will be based on the weighted components above.

PLAGIARISM

The University defines plagiarism in its rules: "Plagiarism involves using the work of another person and presenting it as one's own." Plagiarism is a serious breach of the University's rules and carries significant penalties. You must read the University's practices and procedures on plagiarism. These can be found in the *Handbook of Undergraduate Studies* or on the web at: http://www.student.mq.edu.au/plagiarism/

UNIVERSITY POLICY ON GRADING

Academic Senate has a set of guidelines on the distribution of grades across the range from fail to high distinction. Your final result will include one of these grades plus a standardised numerical grade (SNG).

On occasion your raw mark for a unit (i.e., the total of your marks for each assessment item) may not be the same as the SNG which you receive. Under the Senate guidelines, results may be scaled to ensure that there is a degree of comparability across the university, so that units with the same past performances of their students should achieve similar results.

It is important that you realise that the policy does not require that a minimum number of students are to be failed in any unit. In fact it does something like the opposite, in requiring examiners to explain their actions if more than 20% of students fail in a unit.

The process of scaling does not change the order of marks among students. A student who receives a higher raw mark than another will also receive a higher final scaled mark.

For an explanation of the policy see

http://senate.mq.edu.au/rules/Guidelines2003.doc or
http://senate.mq.edu.au/rules/detailedguidelines.doc.

Obtaining User Account in these labs, each student will be given a user name and password for these labs once they are listed as enrolled in MIST812. After the first time logging into the server, the students need to change their password. The new (changed) password will expire in 30 days and needs to be changed again. If you do not change your password, you will not be able to login to the server again. If this happens, please talk to your tutor or the computer lab administrator:

Mr. Alfred Wong, awong@efs.mq.edu.au phone: 9850 6469

A time-table for the classes scheduled for each week will be displayed on the door of E4B 308. If there is a class in progress, students who are not enrolled in that class are not allowed to use the computers in the lab without permission from the tutor.

Problems with lab computers

Problems with lab computers should be reported as follows:
1. if the problem occurs during a class report problem to your tutor
2. if problem occurs outside class time, then report problem by phone or e-mail to the lab administrator Mr. Alfred Wong awong@efs.mq.edu.au (ext 6469)

USING YOUR MU E-MAIL BROWSER ACCOUNT and no other (staff are instructed to ignore e-mails from Hotmail accounts, etc). BE SURE TO INCLUDE YOUR NAME AND CLASS, THE LAB AND PC NUMBER AND A BRIEF DESCRIPTION OF THE PROBLEM or by using Blackboard send an e-mail to Ayse Bilgin.

Macquarie University provides a range of Academic Student Support Services. Details of these services can be accessed at http://www.student.mq.edu.au.

Other information for Macquarie students is available from the Student Portal https://my.mq.edu.au. This includes access to your official university email.

IT help is available from the IT help desk in the library, and from http://www.lib.mq.edu.au/help/

As a division we also have a resource and information centre for our students, which is located on E4B106.

CONFIDENTIALITY AND COPYRIGHT

The University respects your right to confidentiality in the 'classroom' and the right to have your ideas protected. As such, we aim to develop a safe and secure online learning environment in which you can freely express your ideas and opinions.

Copyright protects the manifestation of your ideas, your words and what you say. In accordance with University policy, as creators of copyright material you have the right:

- to be identified as the creator of a work, and to prevent others from claiming to be the creator of a work (the Right of Attribution)
- the right to not have authorship falsely attributed (the Right against False Attribution)
- the right not to have works subjected to derogatory treatment (the Right of Integrity).

You should be aware that as a routine part of the operations of the Macquarie University Online Teaching Facility automatically collects two types of information:

• Server Usage Information - which contains access information for example, who accesses the site, the location from which the site is accessed, the time and duration of access, and what material is accessed.

• General Unit Management Information - the Blackboard software used in your unit tracks information about how you use the online facilities and teaching materials, for example, your first and last log in, discussion messages you have posted and read, and whether you have read your mail. It is important to note that your convenor cannot read your personal mail or annotations.

More information on confidentiality is available on the Macquarie University Online teaching Facility web site at http://online.mq.edu.au/uw/conf.html

MIST812 Decision Support Systems
UNIT SCHEDULE

WEEK	LECTURE TOPIC	Recommended reading	Assessment Due
W1	Data Mining: Principles and Concepts	Chap1, Han & Kamber 2006	
W2	Data Mining: Data Warehouse and OLAP Technology	Chap 3, Han & Kamber 2006	Lab Ex 1
W3	Data Mining: Data Preprocessing (Missing values, noisy data, inconsistent data) Concept hierarchies, interestingness measures & visualisation	Chap 2 and Chap 4 Han & Kamber 2006	Lab Ex 2
W4	Data Mining: Descriptive data mining	Chap5, Han & Kamber 2001	Lab Ex 3
W5	Data Mining: Classification and Prediction	Chap 6, Han & Kamber 2006	Lab Ex 4 **DM Project 1**
W6	Data Mining: Mining Association Rules - Market Basket Analysis	Chap 5, Han & Kamber 2006	Lab Ex 5
W7	Data Mining: Cluster Analysis	Chap 7, Han & Kamber	Lab Ex 6
SEMESTER BREAK: 20 September – 5 October			
W8	**Public Holiday** **(no lectures or tutorials)**		
W9	GIS: Introduction – presenting business geographic data		Lab Ex 7 **DM Project 2**
W10	GIS: Active earth maps		
W11	GIS: Maps with hyperlinks		
W12	GIS: Contour maps		
W13	Practice Exam Paper will be discussed during the lecture. Tutorial time will be used for presentations.		**GIS project**

Note that all DM lab exercises are due by 6pm