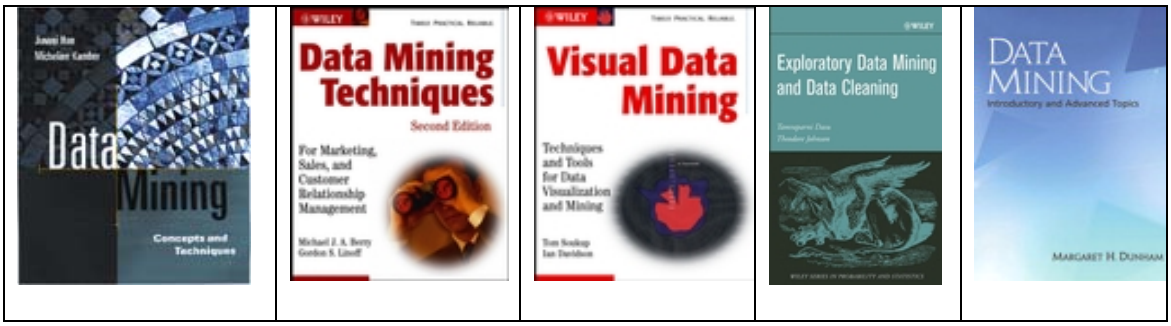
MACQUARIE UNIVERSITY~SYDNEY

# MIST812 – Decision Support Systems
# STAT820 – Decision Support Systems

## Unit Outline
## *Second Semester 2006*




Unit Convenor: 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 one of 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 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/Stat820 is a four credit point unit offered by the Statistics department in the Division of Economic and Financial Studies. This unit expands on topics covered in MIST800 *Computer Applications in Business and a*vailable to graduate students both in Business and in Applied Statistics. It is offered as part of the Master of Commerce/Master of International Business Program.

**Software:**
SPSS, SPSS-CLEMENTINE, MAPINFO, Microsoft Excel & Access

**Prerequisite:**
Basic quantitative knowledge (STAT170 or MIST800 or equivalent).

## TEACHING STAFF

| | | |
|---|---|---|
|  |  | |
| Dr Ayse Bilgin<br>Room: E4A 515<br>Phone: 9850 8509<br>e-mail: abilgin@efs.mq.edu.au | Emeritus Professor<br>Don McNeil<br>Room: E4A 548B<br>Phone: 9850 6473<br>e-mail:<br>dmcneil@efs.mq.edu.au | |

## RECOMMENDED TEXT BOOK

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

## RECOMMENDED REFERENCE TEXTS

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) Data mining introductory and advanced topics by Margaret Dunham, 2003, Prentice Hall/Pearson Education (library call number: QA76.9.D343.D86 2003)

## WEEKLY READINGS

The weekly readings are available through library e-reserve and can be accessed at http://www.library.mq.edu.au/reserve/ . The readings are examinable.

## CLASSES

### Lectures

Lectures begin in Week 1. Students should attend **ONE** 2-hour session per week: Mondays between 6:00 and 8:00pm in E5A131.

### 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 which develops a student thought process and which encourages discussion between students.

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/units/stat820/

## WebCT Access

There is a WebCT site for this subject. All students (MIST or STAT) should log on to STAT820 WebCT site. Students are now required to log into WebCT using their Student ID Number and myMQ Portal Password.

The Web site for WebCT log in is: http://online.mq.edu.au

## Learning outcomes

By the end of this unit students will be able to:

- have an understanding of the principles and the concepts of the data mining
- understand and 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 output 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
- create contour maps and Voronoi polygons
- to understand the link between data mining, geographical information systems and good decision making

## Generic Skills

University study aims, not only to provide you with knowledge and skills in a particular academic discipline, but also to equip you with some generic skills. By studying this unit students will:

- improve their ability to work co-operatively as a team member
- enhance their problem solving skills
- improve their written communication skills, particularly report writing skills
- enhance their critical thinking skills
- be confident in the use of a variety of software packages for solving problems

## Teaching and Learning Strategy

- students are expected to attend all the lectures and the tutorials
- in data mining section, weekly practical exercises are set for individual assessment of lab tasks. However, students can work in groups, but the submitted work must be the students' own work
- the GIS group project will be prepared as a group and presented by every member of the group

- if for any reason, students can not hand in their assessment tasks on time, they have to contact one of the teaching staff in advance
- students should hand in and collect their marked papers from ERIC (Economic Resource & Information Centre) E4B106

## 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/services/policies_consid.htm.
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

**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 hand them into the appropriate box in ERIC (E4B106). These weekly lab exercises will be corrected by your tutor. There will not be any marks for these weekly exercises. It is important to collect them from ERIC so that you can learn from any mistakes you make.

**Data mining projects** will be done in a group of four students. There will be two projects, each worth 10% of total assessment marks. If you want to change your group, you have to find a student from the group that you want to move into who would swap with you. If you can not find a student to swap places, you have to remain in the group that you have been allocated to. If you could find a student who is willing to swap with you please inform Ayse Bilgin.

**GIS project** will be worth 10% of total assessment marks. You will need to work in a group. GIS project is due in week 7. Presentation time table will be available from WebCT.

**Mid semester test (15%)** will be held in the first 40 minutes of the week 6 lecture. This test covers the first five weeks of lecture material and readings. Students may bring one A4 sized sheet of **hand-written** notes, formulas, etc., which may be written on both sides. This summary must be submitted with your test paper. The rest of the lecture will cover new work.

**Final examination (55%)** is 2 hours 30 minutes 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.

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

## 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/

## OVERALL ASSESSMENT

Students are expected to gain a reasonable level of proficiency in weekly topics.

The overall assessment for MIST812 – STAT820 is:

| | |
|---|---|
| Data mining projects (each 10%) | 20% |
| GIS Group Project | 10% |
| Mid semester test | 15% |
| Final examination | 55% |

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

## 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).

**The grades and what they mean are given as below:**

| | |
|---|---|
| **HD - High Distinction** | Denotes a performance that meets all unit objectives in such an exceptional way and with such marked excellence that it deserves the highest level of recognition. |
| **D - Distinction** | Denotes performance that clearly deserves a very high level of recognition as an excellent achievement in the unit. |
| **C -Credit** | Denotes performance that is substantially better than would normally be expected of competent students in the unit. |
| **P - Pass** | Denotes performance that satisfies unit objectives. |
| **PC - Conceded Pass** | Denotes performance that meets unit objectives only marginally. |
| **F - Fail** | Denotes that a candidate has failed to complete a unit satisfactorily. |

For further explanation of the policy see
http://www.mq.edu.au/senate/MQUonly/Issues/Guidelines2003.doc or
http://www.mq.edu.au/senate/MQUonly/Issues/detailedguidelines.doc.

## STATISTICS COMPUTER LABS AND THEIR CONDITIONS OF USE

Apart from EFS computer labs, we have two statistics labs that students can use during the term: E4B202 and E4B308.

### 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 a MIST812 or STAT820. 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 6138

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 6138)

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 WebCT send an e-mail to Ayse Bilgin.

## STUDENT SUPPORT SERVICES

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 and can be accessed at http://www.efs.mq.edu.au/services/eric.htm

**MIST812/STAT820 Decision Support Systems**
**UNIT SCHEDULE**

| WEEK | LECTURE TOPIC | Recommended reading | Assessment Due |
|---|---|---|---|
| **31 July W1** | Data Mining: Principles and Concepts | Chap1, Han & Kamber | |
| **7 Aug W2** | Data Mining: Data Warehouse and OLAP Technology | Chap2, Han & Kamber | Lab Ex 1 |
| **14 Aug W3** | Data Mining: Data Preprocessing (Missing values, noisy data, inconsistent data)<br>Data Mining: Concept hierarchies, interestingness measures & visualisation | Chap3 and Chap4 Han & Kamber | Lab Ex 2 |
| **21 Aug W4** | GIS: Introduction – presenting business geographic data | | Lab Ex 3 |
| **28 Aug W5** | GIS: Active earth maps | | **DM Project 1** |
| **4 Sep W6** | GIS: Maps with hyperlinks<br>**Mid Semester Test** | | |
| **11 Sep W7** | GIS: Contour maps | | **GIS project** |
| **SEMESTER BREAK: 16 September – 2 October** | | | |
| **2 Oct W8** | **Public Holiday (no lectures or tutorials)** | | |
| **9 Oct W9** | Data Mining: Descriptive data mining | Chap5, Han & Kamber | |
| **16 Oct W10** | Data Mining: Mining Association Rules - Market Basket Analysis | Chap6, Han & Kamber | Lab Ex 9 |
| **23 Oct W11** | Data Mining: Classification and Prediction | Chap7, Han & Kamber | Lab Ex 10 |
| **30 Oct W12** | Data Mining: Cluster Analysis | Chap8, Han & Kamber | Lab Ex 11 |
| **6 Nov W13** | Practice Exam Paper will be discussed during the lecture. There are no tutorials. | | Lab Ex 12<br>**DM Project 2** |

***Note that all DM lab exercises are due by 6pm in ERIC E4B106***