

THE UNIVERSITY OF HONG KONG
FACULTY OF BUSINESS AND ECONOMICS
School of Business
Course Syllabus and Outline for Semester I of 2009-10
BUSI0092 Advanced Database Management and Data Mining
Course Web: <http://www.fbe.hku.hk/~bose/adb>

I. INFORMATION ON INSTRUCTOR AND TUTOR

Instructor: Dr. Indranil Bose
Email: bose@business.hku.hk
Office Location: Room 615 Meng Wah Complex
Office Telephone: 2241-5845
Personal Web: <http://www.fbe.hku.hk/~bose/bose.htm>

Instructor for Tutorial Session: Ms. Becky Tsang
Email: beckytyp@hku.hk
Office Location: Room 724 Meng Wah Complex
Office Telephone: 2241-5165

II. COURSE INFORMATION

Prerequisite

BUSI0052 Database Development and Management or
CSIS0278 Introduction to Database Management Systems

Course Description

At the present time organizations can access a large quantity of data about themselves and their environment. Data mining is devoted to unleashing meaningful patterns from voluminous data that can give competitive advantage to businesses. This course provides a three-step overview of the various important concepts that make up the science of data mining. First, it discusses the importance of data warehouses for storage, update, and retrieval of data and their use in modern businesses. Second, it delves into the discussion of data using techniques such as decision trees, neural networks, and clustering through state-of-the-art software. Third, it uses case studies to focus on understanding and evaluating the results of data mining and converting them into recommendable actions for businesses that are aligned with their corporate strategies.

Required Text

Data Mining: A Tutorial Based Primer (RG)
R. J. Roiger and M. Geatz
Addison Wesley
ISBN: 0201741288

Course Delivery

D1. Three hours of class meetings per week that includes lectures, discussions, and problem solving.

D2. One hour of tutorial per week where conceptual and programming problems related to design and management of data mining systems are discussed as well as hands-on learning of data mining software is achieved.

Course Objectives

O1. To learn the theories of data warehousing and data mining.

O2. To learn the different data mining techniques and apply them for data analysis.

O3. To learn to use state-of-the-art data mining software for business data analysis.

O4: To work in teams effectively.

O5. To analyze latest research related to the use of data mining in modern businesses.

Course Learning Goals

G1. To understand how data warehouses differ from databases and how they benefit businesses.

G2. To understand the need for data mining and the various processes that are involved in data mining.

G3. To learn techniques that can be used for mining business data and how to evaluate those techniques.

G4. To use commercial data mining software to analyze real-life business data and draw meaningful conclusions from the data.

G5. To understand the strategies to be adopted by modern businesses for successful use of data mining.

G6. To work in teams.

G7. To learn how to identify major business issues, find relevant information, outline credible alternatives, and logically build and articulate convincing arguments in support of a particular course of action.

G8. To conduct independent research on the latest developments in the world of data mining and determine how they can be applied in the operation of modern businesses.

Course Learning Outcomes

Upon the successful completion of this course, students should be able to

C1. Distinguish between database management system and data warehouse, the requirements for building a data warehouse, and the use of data warehouse in modern businesses.

C2. Understand the conceptual steps involved in data mining and learn how to apply it for solving business problems.

- C3. Evaluate the advantages, disadvantages, and applicability of data mining techniques like decision trees, neural networks, and clustering and learn to use these techniques appropriately for business data mining.
- C4. Use various available functionalities and analysis capabilities of state-of-the-art software to analyze business data and perform hands-on data mining.
- C5. Understand the strategies for the use of data mining in the business world and how it can give competitive advantage to organizations.
- C6. Work in teams effectively.
- C7. Write effective research reports, and make compelling presentations.
- C8. Review, and synthesize existing literature related to cutting edge technological developments in the field of database management from a technical and business perspective.

Alignment of program learning outcomes with course learning outcomes

Program Learning Outcome	Course Learning Outcome
P1. Acquisition and internalization of knowledge of business and information systems	C1, C2, C3, C5, C8
P2. Application and integration of knowledge of business and information systems	C2, C3, C4, C5
P3. Inculcating professionalism and instilling leadership skills	C6
P4. Developing global outlook	N/A
P5. Mastering communication skills	C7

III. TEACHING AND LEARNING ACTIVITIES

TL1. Case Analysis and Presentation

Business cases are used in this course to present a global view of the subject and to challenge you to use your subject relevant knowledge and judgment for effective business decision making in the context of database management and data mining as a team. You will be expected to work in a group, analyze a case, and make a formal presentation of the case in the later part of the semester. Details regarding the formal case presentation and preparation of the case report will be distributed to you later. The presentation is expected to last approximately 20 minutes. Please make sure that all members of the case presenting group are present on the day of their case presentation. You are also required to submit a case report to the instructor. No late submissions will be accepted.

Your grade for the presentation will depend on the ability of your team to deliver a professional and persuasive presentation that identifies key issues related to the case. Your grade for the case report will depend on the following:

- i) identification of major issues and problems in the case,
- ii) application of appropriate concepts to analyze the situation under study,
- iii) a well thought-out and credible recommended course of action, and
- iv) a doable implementation plan.

(Major focus: C1, C2, C3, C5, C7)

TL2. Project and Presentation

An important aspect of the course is to complete a project that allows you to work as a team and enables you to carry out in-depth research on a chosen topic. You will be asked to work in small groups. There will be a project presentation and a report due at the end of the semester. I expect all members of the group to participate equally in all group activities. Any problems with group dynamics need to be resolved as soon as possible. The exact details of what to do for the project will be discussed in details later on.

In your project you are expected to review latest research on a state-of-the-art technology that is related to the theme of this class and present your recommendations on the usability, impact, and relationship of this technological development to the database management and data mining practices of modern businesses. Your grade for the project will depend on the following:

- i) innovativeness of your project,
- ii) completeness of coverage of the topic in your project,
- iii) relevance of your project to the business world, and
- iv) ability to work as a team.

(Major focus: C2, C3, C5, C6, C7, C8)

TL3. Attendance and Participation

The success of this course depends on your individual and collective contribution to the class, tutorial, and discussions. I expect highest level of participation during the entire semester. Attendance is compulsory during regular classes and tutorial sessions and if you miss any class or tutorial session you will be responsible for all material that was discussed. I expect you to read the assigned readings and prepare yourself before each class so that you can contribute effectively to class discussions. You are also expected to remain attentive and alert during the lectures and discussions. Any disruptive behavior in the class will lead to a reduction in class participation points.

Your grade for class participation will depend on your regular attendance of all relevant sessions, ability to listen carefully to all class lectures and discussions, and organize and articulate your questions and remarks, in a logical manner.

(Major focus: C1, C2, C4)

TL4. Assignments

Assignments are an integral part of this course because they test your knowledge of relevant theories and your ability to apply that knowledge for solving practical problems. All assignments are to be individual submissions. The assignments will include conceptual questions, hands-on data analysis exercises using software, and analysis of business cases related to data mining. The assignments are due at the beginning of the designated class day. No fax or e-mail submissions will be accepted. The due date will be strictly enforced and no late submissions will be accepted. No photocopied assignments will be accepted.

Your grade for the assignments will depend on the correctness of your answers to problems, clarity of presentation, and completeness of your solutions.

(Major focus: C1, C2, C3, C4, C5)

TL5. Exam

The exam will be closed book and closed notes and it will be organized in-class. There will be no make-up exam. Conflicts for the exam must be resolved before the exam date. You should contact me at least two weeks prior to the exam date and let me know in writing. Last minute requests will not be entertained. The exam will be comprehensive in nature and it will involve multiple choice questions, short answer type questions, problem solving questions, situational decision making questions or any combination of the above. There will be no make-up exam under any circumstances.

Your grade for the exam will depend on your ability to express your understanding of the theories of database management and data mining, apply the learned theories for solving related problems, and clarity and correctness of your writing.

(Major focus: C1, C2, C3, C5)

IV. PROCESS FOR EVALUATION

Assessment	Points
<u>Group Work (50%)</u>	
TL1. Case Analysis and Presentation	20
TL2. Project and Presentation	30
<u>Individual Work (50%)</u>	
TL3. Attendance and Participation in Lectures and Tutorials	15
TL4. Assignments	10
TL5. Exam	25
TOTAL POINTS	100

V. COURSE POLICIES

Discipline in Class

I will expect you to maintain the decorum of the class at all times. If you have something important to discuss with your friend please leave the class and do so and then come back to the class. Do not whisper in class. If you are late for any class, please enter the class as quietly as you can and occupy any open seat. If you are dissatisfied about grades please come to my office and I will be happy to discuss the issue with you. I will not discuss grades by e-mail. Do not read newspapers and material from other classes during my lecture time. Also please turn off your mobile phones during class time.

Communication

If you need to set up an appointment please send me e-mail. I will respond as quickly as I can. I will also set up a mailing list and make announcements on the list from time to time. You are responsible for checking your e-mail periodically.

Academic Dishonesty

The University Regulations on academic dishonesty will be strictly enforced. Please check the University Statement on plagiarism at <http://www.hku.hk/plagiarism/>

Where a candidate for a degree or other award uses the work of another person or persons without due acknowledgement:

- a. The relevant Board of Examiners may impose a penalty in relation to the seriousness of the offence;
- b. The relevant Board of Examiners may report the candidate to the Senate, where there is prima facie evidence of an intention to deceive and where sanctions beyond those in (a) might be invoked.

VI. COURSE OUTLINE (tentative and subject to change)

Date	Topic	Readings and Deliverables
Sept 1	General Course Introduction	
Sept 3	Review of Relational DBMS	
Sept 8	Data Warehousing	
Sept 10	Data Warehousing	
Sept 15	Data Warehousing at Wal-Mart	Group Names for Project Due
Sept 17	Introduction to Data Mining	
Sept 22	Introduction to Data Mining	RG Chaps 1, 2, 5
Sept 24	Introduction to Case Analysis Project and Case Allocation	
Sept 29	Guest seminar: Business Intelligence Speaker: Mr. Alex Hung Managing Director, Crossover International	RG Chaps 1, 2, 5
Oct 1	No Class	
Oct 6	Decision Trees	RG Chap 3 (3.1, 3.2, 3.5)
Oct 8	More on Decision Trees	RG Chap 3 (3.1, 3.2, 3.5) Assignment 1 Due
Oct 13	No class - Readings Week	
Oct 15	No class - Readings Week	
Oct 20	Guest Seminar: Business Data Mining Speaker: TBA	Proposal for Project Due
Oct 22	Case 1: Bold New Statements from Natwest	

Oct 27	Neural Networks	RG Chap 8
Oct 29	Case 2: A Fruitful Passion for Orange	
Nov 3	Clustering	RG Chap 3 (3.3) Assignment 2 Due
Nov 5	Case 3: Harrah's Entertainment Inc.	
Nov 10	Web Data Mining	
Nov 12	Case 4: Google Inc.: Launching Gmail	
Nov 17	Exam	
Nov 19	Review for Project	
Nov 24	Project Presentations	
Nov 26	Project Presentations	Project Reports Due