CSCI527
ADVANCED DATABASES & DATA MINING
Fall, 2012
Tentative Course Syllabus

- **INSTRUCTOR:**
  Sang C. Suh, Ph.D.
  Professor & Head, Department of Computer Science
  Texas A&M University - Commerce
  Office Hours: Tue (1:00PM-2:00PM), Wed (1:00AM-2:00PM, 4-5PM) & Thur (3:00-6:00PM)
  Office: JOUR 122; Phone: 903.468.8199; E-mail:sang.suh@tamuc.edu

- **CLASS MEETINGS:**
  001 80725  R 1:00-3:40PM  JOUR 234  Suh S

- **COURSE OBJECTIVES:**
The course objectives are two-fold:
  - To learn general theory, concept, and techniques related to the design of intelligent databases,
  - To practice the design of an intelligent database by performing a design project.

- **STUDENT LEARNING OUTCOMES (SLO):**
  1. Understand current status of the state-of-the-art data mining methodology in industry and academics
  2. Obtain the technique for team play and teamwork for large intelligent database projects through brain storming and joint requirement planning
  3. Learn and use effective tools for web navigation and program integration management
  4. Identify dirty data sources and construct data cleaning programs
  5. Construct programs for capturing association rules
  6. Write programs for trend analysis using statistical data mining techniques
  7. Implement code for generating decision rules using decision tree based classification
  8. Apply divide-and-conquer approach and learn to integrate various programs of small size to form a solution to a large integrated program
  9. Learn to apply various data mining techniques into various areas of different domains
  10. Learn how to design a large scale software analysis and design project with a focus on business intelligence
11. Be able to demo and present the initial, intermediate, and final delivery of the system following CMM and rapid prototyping approaches

- **COURSE DESCRIPTION (OUTLINE):**
  
  General theory, concept, and techniques related to intelligent database design are discussed in this course. Many programming techniques to allow students to integrate intelligent database systems with web-based applications are covered in detail. Additional topics to be covered include expert systems, neural networks, hypermedia, and text retrieval. These additional topics are used as supplementary entities to be integrated into database reasoning components. A moderate-size semester project needs to be developed and demonstrated for the practice of the design of an intelligent database. This course project is a purely web-based and practical database system that is capable of handling high-level intelligent queries.

- **TEXTBOOK:**
  

- **RECOMMENDED BOOKS:**
  
  - Gordon S. Linoff and Michael J. A. Berry, Mining the Web, Wiley (2001).

- **TENTATIVE SCHEDULE:**
  
  1. Chapter 1: Introduction (1 Week)
  2. Chapter 2: Association Rules (1 Week)
  3. Chapter 3: Classification Rules (2 Weeks)
  4. Chapter 4: Statistical Data Mining (2 Weeks)
  5. Chapter 5: Rough Sets and Data Mining (2 Weeks)
  6. Chapter 6: Neural Networks in Data Mining (2 Weeks)
  7. Chapter 7: Clustering (1 Week)
  8. Chapter 8: Fuzzy Information Retrieval (1 Week)

- **SOME TOPICS TO BE SELECTED FOR DISCUSSION:**
  
  - Statistical Data Mining
  - Classification
  - Association Rules
  - Clustering
  - Web Mining
  - Fuzzy Logic
  - Rough Set Theory
  - Neural Networks
COURSE REQUIREMENTS:
CSCI526 or equivalent is a prerequisite for this course. It is expected that each student be familiar with relational DBMS including relational data models, ER models, normalization, and functional dependencies. Though each of these will be reviewed in class during the first three weeks of the course, it is very imperative that a basic understanding of them exists. During the first half of the semester, theoretical aspects of relational and intelligent databases will be covered. Most of this material will be derived from instructor’s note, published papers, reference books, and class discussions. Therefore, significant amount of research is expected to be performed by each student to develop appropriate ideas and concepts relevant to intelligent database design. The second half of the semester will be used to develop an intelligent database system by applying concepts and ideas derived during the first half of the semester. Each development project will be different in that each system should display a level of unique intelligent features identified during the first half of the course. Additional requirements of the course are a number of presentations to be made by each team, a test, a term paper, and a project report. Due dates for all assigned materials will be announced in class in advance. It is the student’s responsibility to have all assignments ready on time. Any student who has to be absent on an assignment due date must arrange to have the assignment submitted early. Late assignment may not be accepted.

STUDENTS WITH DISABILITIES ACT COMPLIANCE:
Students requesting accommodations for disabilities must go through the Academic Support Committee. For more information, please contact the Director of Disability Resources & Services, Halladay Student Services Bldg., Room 303D, (903) 886-5835

ACADEMIC ETHICS:
"All students enrolled at the University shall follow the tenets of common decency and acceptable behavior conducive to a positive learning environment." (See Student's Guide Handbook, Policies and Procedures, Conduct).

ATTENDANCE POLICY:
Students are expected to be present at all class lectures. If a student is absent from class on the due date of any assignment, they are expected to make alternative arrangements to assure that the assignment is turned in ON TIME. Any student wishing to withdraw from the course must do so officially as outlined in the class schedule. THE INSTRUCTOR CANNOT DROP OR WITHDRAW ANY STUDENT.

COURSE REQUIREMENT DEADLINES:
Credit will be given for ONLY those exam(s), program(s), and/or project(s) turned in no later than the deadline(s) as announced by the instructor of this class unless prior arrangement has been made with the instructor.
METHOD OF EVALUATION (TENTATIVE):

One Test* (30%)
Database Development (20%)
Project Portfolio (20%)
Design Project Specification† (20%)
Presentation (10%)

LITERATURE SEARCH AND RESEARCH ASSISTANCE:

For research assistance, please contact:

Reference Librarian: Sciences & Collection Development
Karen_Akins@tamu-commerce.edu
903-886-5728

Students with Disabilities:
The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you have a disability requiring an accommodation, please contact:
Office of Student Disability Resources and Services
Texas A&M University-Commerce
Halladay Student Services Building
Room 303 A/D
Phone (903) 886-5150 or (903) 886-5835
Fax (903) 468-8148
StudentDisabilityServices@tamu-commerce.edu

* The one test may be replaced by the project.
† Each design project can be performed by a team of several students.