CSCI 6350 : Data Mining
Spring 2006

1. Instructor

Dr. Cen Li

Office:	KOM 361A
Phone:	(615) 904-8168
E-mail:	**cli@mtsu.edu**
URL:	http://www.mtsu.edu/~cli
FAX:	(615) 898-5567

Office Hours:	Tuesday 9:30 – 10:30 am
	Thursday 1:30 – 3:00 pm
	Others by appointment

2. Course Description:

Prerequisite: CSCI 3110 and CSCI 4350/5350.

Course Objectives: This course introduces students to the fundamental concepts, ideas, techniques and applications of the field of data mining. Techniques such as scientific discovery/modeling, association rule analysis, classification, clustering, deviation detection, statistical modeling approaches, and the emergent technologies are discussed. Students will be taught to work with a number of data mining software on real world data sets to uncover interesting and useful patterns. A final project is required which may be in the form of a reimplementation of a data mining algorithm, development of a new data mining software, or an application of multiple data mining techniques on analyzing a real world data of interest. A comprehensive report describing the project, including analysis of experimental results is required.

3. Learning Outcomes:

o   Students will understand the data mining problems arise from various domains, including scientific, market transaction, warehousing, bio-medical, and the web.
o   Students will understand the basic data mining techniques including association rule analysis, classification analysis, clustering analysis, text mining, trend analysis, and various statistical modeling approaches.
o   Students will be able to apply association rule analysis, classification analysis, clustering analysis, text mining and statistical modeling on real world data to discover interesting and potentially useful patterns.
o   Students will understand the basic techniques used in developing a web search engines and understand how to apply data mining techniques in improving the efficiency of web search engines and in improving the web user satisfiability.
o   Students will understand the basic techniques used in statistical modeling involving temporal data and other structured data.

4. Course Materials:

Text books:
    J.W. Han and M. Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers.
    Additional readings from Journals and conference proceedings will be distributed in class.

References:
    I.H. Witten and E. Frank, "Data Mining: Practical Machine, Learning Tools and Techniques with JAVA Implementations", Morgan Kaufmann Publishers.

5. Methods of Instructions:

    The format of the class will mainly be lecture-based. There will also be classes devoted to paper discussion. Active participation of paper discussion is encouraged.

    Programming assignments and homework assignments need to be worked on individually. Each assignment will be assigned a due date. The assignments are due at the beginning of class time on the due date. Late penalty is 10% deduction per class day. No late homework will be accepted after 2 class days.

    Students are required to select a term project to work on. This includes survey of the background material, written project proposal (1-2 pages), actual implementation and testing, and class presentation. Term projects can be worked on individually, except for when the project involves extensive programming, in which case, a group of two students may work on the same project. The last week of the semester will be devoted to term project presentation.

    There will be two tests and one final exam. All tests will be closed book. The final exam is comprehensive. No makeup test will be given.

6. Grading Policy:

Your grade will be determined as following:
    25 %   Assigned projects and homework
    25 %   Term project(s) and presentation
    30 %   Tests
    20 %   Final exam

Letter grades will be assigned strictly according to the following scale:
    A   average at least 90%
    B+ average at least 88% but under 90%
    B   average at least 82% but under 88%
    B- average at least 80% but under 82%
    C+ average at least 78% but under 80%
    C   average at least 72% but under 78%
    C- average at least 70% but under 72%
    D+ average at least 68% but under 70%
    D   average at least 62% but under 68%

      D-   average at least 60% but under 62%
      F   average under 60%

7. Attendance

Attendance is required and absences do not excuse one from class responsibilities. If for some unavoidable reason you must miss class, obtain class notes, handouts, and assignments from another class member or from the course page. Generally, you will find handouts and assignments posted at course web page at www.mtsu.edu/~cli.

8. Important Dates:

Last day to drop without a 'w':	January 30th
Last day to drop with a 'w':	March 6th
Spring break (no class):	March $6^{th} – 11^{th}$
Last day of classes:	April 26th
Final exam:	Tuesday, May 2nd, 3:30 – 5:30 pm

6. Important Notices:

(a) *Disabled Student Notice:*
If you have a disability that may require assistance or accommodation, or you have questions related to any accommodations for testing, note takers, readers, etc., please contact me as soon as possible. Students should also contact the Office of Disabled Student Services (898-2783) with questions about such services.

(b) If you bring a mobile phone or pager to class, it must be either turned off or set to signal you silently.

(c) If you are having problems with the course or have a complaint you would like to voice, please bring this to the attention of the course instructor as soon as possible.