

## COURSE DESCRIPTION

Dept., Number	CSci 581/582	Course Title	Special Topics in Computer Science I/II: Data Mining, Summer 2006
Semester hours	3	Course Coordinator	Dawn E. Wilkins, Associate Professor

### Current Catalog Description

Special topics in computer science.

Topic description: Knowledge discovery in databases (KDD) applies techniques from artificial intelligence, statistics, and pattern recognition to detect patterns in large databases. The goal is to discover previously unknown patterns and cause and effect relationships in the data. Data mining is one step in the KDD process.

### Textbook

Jiawei Han and Micheline Kamber. *Data Mining: Concepts and Technique*, 2<sup>nd</sup> edition, Morgan Kaufmann Publishers, March 2006.

### References

Website: <http://www.cs.olemiss.edu/~dwilkins/mining/sum06/syllabus.html>

### Course Outcomes

Upon successful completion of this course, the students:

1. are aware of the steps in the overall KDD process;
2. understand how to use data analysis techniques to make better business decisions;
3. know when it is appropriate to use the various data preparation and analysis techniques;
4. know the strengths and weaknesses of various data mining tools and methods, including clustering, association rule creation, and prediction (neural networks and decision trees);
5. can use data mining tools for building models for various application domains.

### Relationship between Course Outcomes and Program Outcomes

1. Be aware of the steps in the overall KDD process. Outcome: (b)
2. Understand how to use data analysis techniques to make better business decisions. Outcome: (j)

3. Know when it is appropriate to use the various data preparation and analysis techniques. Outcome: (b), (c) and (j)
4. Know the strengths and weaknesses of various data mining tools and methods, including clustering, association rule creation, and prediction (neural networks and decision trees). Outcome: (b)
5. Demonstrate the use of data mining tools for building models for various application domains. Outcome: (b), (c) and (j).

#### Prerequisites by Topic

The official prerequisites for CSci 581 and 582 are CSci 211 (Computer Science III) and CSci 223 (Computer Organization and Assembly Language). However, the students in this section should also have an interest in databases and/or artificial intelligence and a familiarity with basic statistics.

#### Major Topics Covered in the Course

Introduction/review of database and data warehousing, multidimensional data model, data mining tasks, steps in KDD process, data preprocessing (cleaning, integration, transformation, reduction, discretization), concept descriptions, association rules, classification and prediction (decision tree induction and neural networks), clustering, mining complex data (spatial databases, multimedia databases, time-series, text databases, WWW), visualizations, applications.

#### Assessment Plan for the Course

This is an elective course offered on this topic only once with the specified content. The offering had 4 examinations and 3 homework assignments. The course outcomes were assessed by the examinations and assignments.

How Data in the Course are Used to Assess Program Outcomes (unless adequately covered already in the assessment discussion under Criterion 4)

--

#### Estimate Curriculum Category Content (Semester hours)

Area	Core	Advanced	Area	Core	Advanced
Algorithms		2	Software design		
Data structures		1	Concepts of programming languages		