ISC4933/5935   Data Mining

Fall Semester 2014

Day / Time: MWF 09:05 AM – 09:55 AM

Location: DSL 0499

Instructor: Dr. Anke Meyer-Baese
Email: ameyerbaese@fsu.edu
Office: Room 476 DSL Dirac Science Library
Office Hours: 10:00 – 12:00 PM M or by appointment
Phone: 644-3494

Textbook: **Introduction to Data Mining**, by P. N. Tan et al., Pearson, 2006.

Course Web Page: A course web page is being developed and will have a hyperlink on http://campus.fsu.edu/. An announcement will be made in class when the web page is completed.

Catalog Description: Basic data mining concepts – data representation and visualization. Classification techniques: decision trees, rule-based classifier, nearest-neighbor classifier, Bayesian classifier, artificial neural networks, support vector machines. Cluster analysis: density-based cluster, graph-based cluster. Basic learning mechanisms: supervised and unsupervised. Temporal and spatial mining: prediction, time-series, regression. Performance evaluation: ROC curves, confusion matrix. Applications of data mining: anomaly detection, remote sensing, bioinformatics and medical imaging. Programming exercises will be assigned.

Prerequisites: ISC3222 or ISC3313 or ISC4304 or COP 3330 or consent of instructor.

Course Goals: At the completion of this course the student should know:

1. Basic data mining tasks and metrics.

2. Data mining techniques.

3. Classification.

4. Clustering.

5. Spatial and temporal mining.

6. Evaluation techniques.

7. Application of data mining: anomaly detection, bioinformatics and medical imaging.

Class Policies:
Exams/Tests: - Test dates announced at least 1 week in advance.
- Quizzes will be given without notice.
- No make-ups will be granted unless **prior** approval has been obtained from the instructor.

Homework:	- Assignments are due at the *BEGINNING* of class on the due date.
	- Late assignments will be assessed a *50% penalty* for the *first 24 hours.*
	- Assignments will *NOT BE ACCEPTED MORE THAN 24 HOURS LATE.*

Attendance: - **Class attendance is required for all students.** College and University rules allow only 3 unexcused absences for this course. A student **Exceeding 3 unexcused absences** will be **dropped from the course and assigned a grade of "F".**
- In-class assignments will be the primary method for taking attendance.

Ethics/Honor Code: - All students are bound by the honor code of their university. Violations of the honor code will be reported. Penalties include but are not limited to 1) failing grade on the assignment and 2) failing grade for the course.
- All assignments are considered *individual* efforts. Students are encouraged to discuss topics and homework, but the work itself is to be performed on an individual basis.

Grading Policy: ISC4933/5935 :

Mid-term test	40%
Term research project	40%
In-class assignments	10%
Homework Assignments	10%

Questions, problems and errors involving the grading of any assignment or test must be brought to the attention of the instructor **within 1 week** of the graded work's return to the *class*. A student's absence from class does not extend the time limit. After 1 week the grade is final and will not be reviewed at the student's request.

Students with Disabilities: Students with disabilities needing academic accommodations should: (1) Register with and provide documentation to the Student Disability Resource Center (SDRC); and (2) Bring a letter to the instructor indicating the need for accommodations and what type. This should be done within the first week of class. *This syllabus and other class materials are available in alternative format upon request*.

Tentative **Class Schedule** (subject to modification; provided only for planning purposes)

Week	Topic
1	Introduction to data mining tasks.
2	Data mining metrics.
3	Data mining techniques.
4	Classification, introductory concepts.
5	Classification, decision trees and alternative techniques.
6	Clustering, basic techniques.
7	Review and Midterm Test
8	Clustering, supervised and unsupervised techniques.
9	Clustering, graph-based algorithms.
10	Spatial and temporal mining.
11	Application of data mining: anomaly detection.
12	Application of data mining: bioinformatics
13	Applications of data mining: remote sensing

Note: Dates and material covered are subject to modification at any time.