CSE 8331 - Spring 2012
Advanced Topics in Data Mining

http://michael.hahsler.net/SMU/8331

Introduction

This is the second course in Data Mining. A prerequisite is successful completion of CSE 7331 or another introductory Data Mining course. Please contact Dr. Hahsler if you have concerns or questions about this prerequisite. It is assumed that every student is familiar with basic data mining methods (clustering, classification, and association rules) and has some experience with programming and one or more data mining tools (R, RapidMiner, Weka, XLMiner, etc.).

The objective of this course is to get an overview of several advanced data mining techniques and understand the research methods applied in the field.

In this course we will investigate a set of advanced data mining techniques by:

- Reading, reviewing and discussing research papers.

- Preparing and presenting a tutorial on a specific topic.

- Conducting a research project (applying an advanced data mining method, implementing a method or preparing a review paper).

Instructor Contact Information

Dr. Michael Hahsler
Caruth 451
(214) 768-8878
mhahsler@lyle.smu.edu
Office hours: by appointment

Department Contacts

Debra McDowell, debra@lyle.smu.edu, (214) 768-3080

Beth Minton, beth@lyle.smu.edu, (214) 768-3083

Text

Research papers will be posted on the web page. Textbooks for reference (recommended but not required) are:

1. Data Mining Introductory and Advanced Topics by Margaret H. Dunham, Prentice Hall, 2003.

2. Introduction to Data Mining by Pang-Ning Tan, Michael Steinbach and Vipin Kumar, Addison Wesley, 2005.

3. The Elements of Statistical Learning by Trevor Hastie, Robert Tibshirani and Jerome Friedman, 2nd edition, Springer, 2009.

Course Work and Grading

The final course grade will be based upon performance on various assignments and class participation. The percentage break-down is as follows:

Assignment	Percentage
Tutorial preparation and presentation	40.00%
Research paper review and discussions	20.00%
Final project and presentation	40.00%

All assignments have to be turned in on time. No exceptions are granted. This also applies to distance students. Details about the assignments will be posted on the course web site. Final grades in this course are determined as follows:

93 - 100 A	80 - 82 B-	67 - 69 D+
90 - 92 A-	77 - 79 C+	63 - 66 D
87 - 89 B+	73 - 76 C	60 - 62 D-
83 - 86 B	70 - 72 C-	00 - 59 F

Covered Topics

Session	Date	Lecture	Presenter
Week 1	1/18	**M- No class**, W – Course overview	Dr. Hahsler
Week 2	1/23	M - Introduction to data stream mining, W – Paper discussion	Dr. Hahsler
Week 3	1/30	Individual meetings: discuss topic and project	-
Week 4	2/6	M – Tutorial, W – Paper discussion	Student, Dr. Hahsler
Week 5	2/13	M – Tutorial, W – Paper discussion	Student, Dr. Hahsler
Week 6	2/20	M – Tutorial, W – Paper discussion	Student, Dr. Hahsler
Week 7	2/27	M – Tutorial, W – Paper discussion	Student, Dr. Hahsler
Week 8	3/5	M – Tutorial, W – Paper discussion	Student, Dr. Hahsler
Week 9	3/12	**Spring break**	
Week 10	3/19	Individual meetings: mid-term progress report	-
Week 11	3/26	M – Tutorial, W – Paper discussion	Student, Dr. Hahsler
Week 12	4/2	M – Tutorial, W – Paper discussion	Student, Dr. Hahsler
Week 13	4/9	M – Tutorial, W – Paper discussion	Student, Dr. Hahsler
Week 14	4/16	M – Tutorial, W – Paper discussion	Student, Dr. Hahsler
Week 15	4/23	Final project presentation	Students
Week 16	4/30	**M – Last day of class,** Final project presentation	Students

Learning Outcomes

After successful completion of this course, you should be able to:

1.0 – DEMONSTRATE COMPETENCY IN INDEPENDENT RESEARCH.

    1.1 Know how to find and evaluate relevant technical information on a topic.

    1.3 Know how to read, and evaluate CS research paper on a newly developing topic.

2.0 – DEMONSTRATE COMPETENCY IN ADVANCED DATA MINING TECHNIQUES.

    2.1 Gain an understanding of the problems and techniques used for a set of advanced data mining application areas.

    2.2 Learn the detailed concepts of a specific advanced data mining topic.

2.3 Use/implement/analyze tools to process and analyze data.

Attendance Policy

Because of the nature of this class, attendance of and participation in this course is very important. Therefore, students are expected to attend regularly. If a student is absent from class, it is that student's responsibility to make arrangements with the professor to make up any work missed or to ensure that assignments are submitted on time or early. Late assignments will not be accepted except in extreme instances. Any assignments that will be missed (including those due to university-sanctioned events) **must be completed before the due date.**

Academic Ethics and Collaboration

All submitted work is expected to be your own. In particular this aims at plagiarism (using someone else's words or ideas without proper citation). Here are your options:

- **Direct Quotation.** You need quotation marks and a reference. Use this very sparingly to never.
- **Paraphrasing.** Put the main ideas/arguments of someone else into your own words and supply references which make clear what you got from where. Paraphrasing is difficult since you cannot repeat parts of the original wording in your own work! This is the most frequently used method to talk about someone else's work.
- Everything in your paper that is not clearly marked as someone else's idea/work must be 100% your own!

You will receive an automatic 0 on an assignment where you have been found copying or plagiarizing. In severe cases, you will receive an F in the course.

Additional Information

Disability Accommodations: Students needing academic accommodations for a disability must first be registered with Disability Accommodations & Success Strategies (DASS) to verify the disability and to establish eligibility for accommodations. Students may call 214-768-1470 or visit http://www.smu.edu/alec/dass to begin the process. Once registered, students should then schedule an appointment with the professor to make appropriate arrangements.

Religious Observance: Religiously observant students wishing to be absent on holidays that require missing class should notify their professors in writing at the beginning of the semester, and should discuss with them, in advance, acceptable ways of making up any work missed because of the absence. (See University Policy No. 1.9.)

Excused Absences for University Extracurricular Activities: Students participating in an officially sanctioned, scheduled University extracurricular activity should be given the opportunity to make up class assignments or other graded assignments missed as a result of their participation. It is the responsibility of the student to make arrangements with the instructor prior to any missed scheduled examination or other missed assignment for making up the work. (University Undergraduate Catalogue)