

EMIS/CSE 8331 Advanced Data Mining

Spring 2017

Instructor: Michael Hahsler
Office: 311 Caruth
Phone: (214) 768-8878 (email preferred)
E-mail: mhahsler@lyle.smu.edu
Office hours: Tue 11:00-12:30pm and Thu 9:00-10:30am or by appointment

Course Information

Lecture time: TTh 2:00-3:20pm
Course Website: <http://michael.hahsler.net/SMU/8331>

Introduction

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).

Prerequisites

- EMIS 5/7331 (formerly EMIS 5/7732) or CSE 5/7331 or an introductory course on data mining.
- It is assumed that every student is familiar with basic data mining methods (clustering, classification, and association rule mining) and has some experience with programming and one or more data mining tools (R, RapidMiner, Weka, XLMiner, etc.). Please contact the instructor if you are not sure if you are sufficiently prepared for this course.

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.

Reading

Weekly reading (research papers) will be posted on the web page.

Textbooks for reference (recommended but not required) are:

1. Introduction to Data Mining by Pang-Ning Tan, Michael Steinbach and Vipin Kumar, Addison Wesley, 2005.
2. The Elements of Statistical Learning by Trevor Hastie, Robert Tibshirani and Jerome Friedman, 2nd edition, Springer, 2009.
3. Data Mining Introductory and Advanced Topics by Margaret H. Dunham, Prentice Hall, 2003.

Course Work and Grading

The final course grade will be based upon performance on various assignments and class participation. You can find details about the assignments on the course web site. The percentage break-down is as follows:

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

Covered Topics

Session	Date	Lecture	Presenter
Week 1	1/24 & 26	T – Course overview, Th – Individual meetings: discuss topic	Dr. Hahsler
Week 2	1/31 & 2/2	T – Individual meetings, Th – Paper discussion	Dr. Hahsler
Week 3	2/7 & 9	T – Paper discussion, Th – Tutorial (Data Stream Mining)	Dr. Hahsler
Week 4	2/14 & 16	T – Paper discussion, Th – Tutorial	Student, Dr. Hahsler
Week 5	2/21 & 23	T – Paper discussion, Th – Tutorial	Student, Dr. Hahsler
Week 6	2/28 & 3/2	T – Paper discussion, Th – Tutorial	Student, Dr. Hahsler
Week 7	3/7 & 9	T – Paper discussion, Th – Tutorial	Student, Dr. Hahsler
Week 8	3/14 & 16	Spring break	-
Week 9	3/21 & 23	T – Paper discussion, Th – Tutorial, Progress report due (3/23)	Student, Dr. Hahsler
Week 10	3/28 & 30	T – Paper discussion, Th – Tutorial	Student, Dr. Hahsler
Week 11	4/4 & 4/6	T – Paper discussion, Th – Tutorial	Student, Dr. Hahsler
Week 12	4/11 & 13	T – Paper discussion, Th – Tutorial	Student, Dr. Hahsler
Week 13	4/18 & 20	T – Paper discussion, Th – Tutorial	Student, Dr. Hahsler
Week 14	4/25 & 27	Project presentations, Final report due (midnight, 4/24)	Students
Week 15	5/2 & 4	Project presentations	Students

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 Catalog)