CSE 5334:  Data Mining, Fall 2019

Tuesdays 7 - 9:50pm  Location: ERB 130

Instructor: Dr. Chris Ding, 529 ERB, 204 ERB. Email: chqding@uta.edu
Office Hours: Mon/Tues 2 – 4 pm (and by appointment).

Teaching Assistant: Qicheng Wang qicheng.wang@mavs.uta.edu
Office Hours: Tues/Thur 3– 5 pm, ERB 204

About the Instructor
Prof. Ding obtained Ph.D. from Columbia Univ and worked at Caltech, Jet Propulsion
Lab, Lawrence Berkeley Natl Lab before joining UTA. He is a leading researcher in
machine learning. His work on k-means clustering, nonnegative matrix factorization, L21
matrix norm, feature selection, etc, are well-known. His publications have been cited
41000 times. He has given invited lectures in UC Berkeley, Stanford, CMU, Google
Research, IBM Research, Microsoft Research, etc. He has taught Data Mining for 5 years.

Contents and Objectives:
Data mining (DM) is often defined as knowledge discovery in database (KDD). Today,
DM is a broad area of data analysis, exploration, using techniques from Machine
Learning, Artificial Intelligence, Statistics and Database. This course will cover main
topics, including classification, clustering, association rule discovery, feature selection,
dimension reduction, semi-supervised learning.

After completing this course, students will be able independently analyze data, finding
patterns in it, design and implement practical algorithms to solve classification problems,
such as recognize the hand-written digits/alphabets, and improve the classification tasks
by using feature selection and dimension reductions.

Prerequisites:
Undergrad level (competent) Linear Algebra and Statistics, Computer Algorithms

Textbook:

Textbook:
An Introduction to Data Mining, 2nd Edition (homework refers to 1st Edition)
By Tan, Steinbeck, Karpatne, Kumar

The class will have

-  Homeworoks (10%)
-  3 written exams (40%)
-  3 computer projects/homeworks
- 3 computer exams (50%).
- Computer projects can be done by a team of 1 or 2 students. The codes must be submitted. Computer exam will make use of the codes developed in these projects.

The weights for final grade is given in percentages above. The grading is based on students’ relative performances. Thus whether a particular exam is easy or hard will not change the relative ranking and thus final grade.

The final exam

- will consists of a written exam (60mins) and a computer exam (60mins)
- Date: Tuesday December 3, 2019. Starting 7:10pm. Location ERB 130

Attendance is highly recommended
Because the rapid advance of the field, we cover the most recent advances, which are not all covered in the textbooks.

- Americans With Disabilities Act
The University of Texas at Arlington is on record as being committed to both the spirit and letter of federal equal opportunity legislation; reference Public Law 93112 -- The Rehabilitation Act of 1973 as amended. With the passage of new federal legislation entitled Americans With Disabilities Act - (ADA), pursuant to section 504 of The Rehabilitation Act, there is renewed focus on providing this population with the same opportunities enjoyed by all citizens. As a faculty member, I am required by law to provide "reasonable accommodation" to students with disabilities, so as not to discriminate on the basis of that disability. Student responsibility primarily rests with informing faculty at the beginning of the semester and in providing authorized documentation through designated administrative channels.

- Academic Dishonesty
It is the philosophy of The University of Texas at Arlington that academic dishonesty is a completely unacceptable mode of conduct and will not be tolerated in any form. All persons involved in academic dishonesty will be disciplined in accordance with University regulations and procedures. Discipline may include suspension or expulsion from the University. "Scholastic dishonesty includes but is not limited to cheating, plagiarism, collusion, the submission for credit of any work or materials that are attributable in whole or in part to another person, taking an examination for another person, any act designed to give unfair advantage to a student or the attempt to commit such acts." (Regents’ Rules and Regulations, Part One, Chapter VI, Section 3, Subsection 3.2, Subdivision 3.22)

Student Support Services Available
The University of Texas at Arlington supports a variety of student success programs to help you connect with the University and achieve academic success. These programs include learning assistance, developmental education, advising and mentoring, admission
and transition, and federally funded programs. Students requiring assistance academically, personally, or socially should contact the Office of Student Success Programs at 817-272-6107 for more information and appropriate referrals.