CSE5301 Data Analysis and Modeling Techniques

http://crystal.uta.edu/~gao/CSE5301_DataAnalysis/Syllabus.html

Instructor: Jean Gao
Email: gao@uta.edu
Office Hour: Tue and Thu, 12:00 - 1:00pm or by appointment.

TA: Fariba Khoshghalbvash
Email: fariba.khoshghalbvash@mavs.uta.edu
Office: ERB 545
Office Hour: Tue & Thu, 3:30pm - 5pm or by appointment

Course Description and Objectives:

The objective of this course is to provide students the basic data analysis and modeling concepts and methodologies using probability theory. Basic statics concepts and probability concepts will be covered. Fundamental data analysis and hypothesis techniques will be covered. Further data modeling methodologies such as Hidden Markov Models and Bayesian networks will be introduced.

Students successfully completing this course will have gained a solid understanding of probabilistic data modeling, interpretation, and analysis and thus have formed an important basis solving practical statistics and data analysis related problems arising in broad computer science and engineering, and daily life.

Prerequisites:

All students are expected to have a background in basic probability, Calculus, and Algebra before attending this course.

Textbook:


References: (under 2 hours reservation in the library):

• Additional course materials will be available electronically through the course website.

Evaluation:

Homework (10%): There will be about 5~6 HWs. Some of the homework assignments may include small computer projects. You are free to choose your most comfortable programming language. Hardcopy of homework assignments are collected before class starts.

Exams (75%): There will be three exams, each will cover approximately one-third of the course materials.

Quizzes (15%): There will be three 15-minute quizzes.

Tentative Major Topics to Be Covered (subjective to change based on course progress):

1. Basic probability: Discrete and continuous random variables, independence, covariance, central limit theorem, Chebyshev inequality, diverse continuous and discrete distributions.

2. Statistics, Parameter Estimation, and Fitting a Distribution: Descriptive statistics, method of moments, maximum likelihood estimation


4. Hypothesis Testing: Type I and II errors, rejection regions; Z-test, T-test, Bayesian test, Confidence Intervals

5. Data Modeling: Hidden Markov Models, Bayesian Network, Regression, Queuing systems

Course Policy:

• Class Attendance is required.

• Cooperative efforts at understanding the material and the assignments are encouraged. However, you are required to present your work that you have completed individually. There will be no make-up exams or quizzes for this course unless the instructor is notified IN ADVANCE under extenuating circumstances.
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 93-112 -- 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.