

**CSE 4334/5334: Data Mining**  
Spring 2019

**Instructor(s):** Won Hwa Kim

**Office Number:** SEIR #324

**Office Telephone Number:** N/A

**Email Address:** won.kim@uta.edu

**Faculty Profile:** <http://ranger.uta.edu/~wonhwa>

**Office Hours:** Mon 11:00am – 12:00pm

**Section Information:** CSE4334 - 001 / CSE5334 - 001

**Time and Place of Class Meetings:** ERB #130, MW 1:00pm – 2:20am

**Description of Course Content:** This is an introductory course on data mining. Data Mining refers to the process of automatic discovery of patterns and knowledge from large data repositories, including databases, data warehouses, Web, document collections, and data streams. We will study the basic topics of data mining, including data preprocessing, frequent pattern mining, correlation analysis, machine learning methods for classification, prediction and clustering.

**Student Learning Outcomes:** A solid understanding of the basic concepts, principles, and techniques in data mining; an ability to analyze real-world applications, to model data mining problems, and to assess different solutions; an ability to design, implement, and evaluate data mining software.

**Required Textbooks and Other Course Materials:** Introduction to Data Mining, 2<sup>nd</sup> Edition, Pearson

**Descriptions of major assignments and examinations:** Written/Programming Assignments, Pop Quiz, and Final exam.

**Attendance:** Attendance is strongly recommended.

**Grading:**

Assignments: 30% (3 assignments expected)

(Pop) Quizzes: 40% (5 to 6 quizzes expected)

Final Exam: 30%

- CSE 4334 and 5334 will be graded separately.
- The grading criteria above is subject to change.
- There is no extra credit materials.
- One lowest score from pop quiz scores will be dropped.
- Regrading policy: Within 7 days after scores are posted, TA will handle regrade requests. If not satisfied with the results, 7 days to request again. Instructor will handle it, and the decision is final.

**Drop Policy:** Students may drop or swap (adding and dropping a class concurrently) classes through self-service in MyMav from the beginning of the registration period through the late registration period. After the late registration period, students must see their academic advisor to drop a class or withdraw. Undeclared students must see an advisor in the University Advising Center. Drops can continue through a point two-thirds of the way through the term or session. It is the student's responsibility to officially withdraw if they do not plan to attend after registering. **Students will not be automatically dropped for non-attendance.** Repayment of certain types of financial aid administered through the University may be required as the result of dropping classes or withdrawing. For more information, contact the Office of Financial Aid and Scholarships (<http://www.uta.edu/aao/fao/>).

**Disability Accommodations:** UT Arlington is on record as being committed to both the spirit and letter of all federal equal opportunity legislation, including *The Americans with Disabilities Act (ADA)*, *The Americans with Disabilities Amendments Act (ADAAA)*, and *Section 504 of the Rehabilitation Act*. All instructors at UT Arlington are required by law to provide “reasonable accommodations” to students with disabilities, so as not to discriminate on the basis of disability. Students are responsible for providing the instructor with official notification in the form of a **letter certified** by the Office for Students with Disabilities (OSD). Only those students who have officially documented a need for an accommodation will have their request honored. Students experiencing a range of conditions (Physical, Learning, Chronic Health, Mental Health, and Sensory) that may cause diminished academic performance or other barriers to learning may seek services and/or accommodations by contacting: **The Office for Students with Disabilities, (OSD)** <http://www.uta.edu/disability/> or calling 817-272-3364. Information regarding diagnostic criteria and policies for obtaining disability-based academic accommodations can be found at [www.uta.edu/disability/](http://www.uta.edu/disability/).

**Counseling and Psychological Services (CAPS)** [www.uta.edu/caps/](http://www.uta.edu/caps/) or calling 817-272-3671 is also available to all students to help increase their understanding of personal issues, address mental and behavioral health problems and make positive changes in their lives.

**Non-Discrimination Policy:** The University of Texas at Arlington does not discriminate on the basis of race, color, national origin, religion, age, gender, sexual orientation, disabilities, genetic information, and/or veteran status in its educational programs or activities it operates. For more information, visit [uta.edu/eos](http://uta.edu/eos).

**Title IX Policy:** The University of Texas at Arlington (“University”) is committed to maintaining a learning and working environment that is free from discrimination based on sex in accordance with Title IX of the Higher Education Amendments of 1972 (Title IX), which prohibits discrimination on the basis of sex in educational programs or activities; Title VII of the Civil Rights Act of 1964 (Title VII), which prohibits sex discrimination in employment; and the Campus Sexual Violence Elimination Act (SaVE Act). Sexual misconduct is a form of sex discrimination and will not be tolerated. *For information regarding Title IX, visit [www.uta.edu/titleIX](http://www.uta.edu/titleIX) or contact Ms. Michelle Willbanks, Title IX Coordinator at (817) 272-4585 or [titleix@uta.edu](mailto:titleix@uta.edu)*

**Academic Integrity:** Students enrolled all UT Arlington courses are expected to adhere to the UT Arlington Honor Code:

*I pledge, on my honor, to uphold UT Arlington’s tradition of academic integrity, a tradition that values hard work and honest effort in the pursuit of academic excellence.*

*I promise that I will submit only work that I personally create or contribute to group collaborations, and I will appropriately reference any work from other sources. I will follow the highest standards of integrity and uphold the spirit of the Honor Code.*

UT Arlington faculty members may employ the Honor Code in their courses by having students acknowledge the honor code as part of an examination or requiring students to incorporate the honor code into any work submitted. Per UT System *Regents’ Rule* 50101, §2.2, suspected violations of university’s standards for academic integrity (including the Honor Code) will be referred to the Office of Student Conduct. Violators will be disciplined in accordance with University policy, which may result in the student’s suspension or expulsion from the University. Additional information is available at <https://www.uta.edu/conduct/>. Faculty are encouraged to discuss plagiarism and share the following library tutorials <http://libguides.uta.edu/copyright/plagiarism> and <http://library.uta.edu/plagiarism/>

**Lab Safety Training:** Students registered for this course must complete all required lab safety training prior to entering the lab and undertaking any activities. Once completed, Lab Safety Training is valid for the remainder of the same academic year (i.e., Fall through Summer II) and must be completed anew in subsequent years. There are no exceptions to this University policy. Failure to

complete the required training will preclude participation in any lab activities, including those for which a grade is assigned.

**Electronic Communication:** UT Arlington has adopted MavMail as its official means to communicate with students about important deadlines and events, as well as to transact university-related business regarding financial aid, tuition, grades, graduation, etc. All students are assigned a MavMail account and are responsible for checking the inbox regularly. There is no additional charge to students for using this account, which remains active even after graduation. Information about activating and using MavMail is available at <http://www.uta.edu/oit/cs/email/mavmail.php>.

**Campus Carry:** Effective August 1, 2016, the Campus Carry law (Senate Bill 11) allows those licensed individuals to carry a concealed handgun in buildings on public university campuses, except in locations the University establishes as prohibited. Under the new law, openly carrying handguns is not allowed on college campuses. For more information, visit <http://www.uta.edu/news/info/campus-carry/>

**Student Feedback Survey:** At the end of each term, students enrolled in face-to-face and online classes categorized as "lecture," "seminar," or "laboratory" are directed to complete an online Student Feedback Survey (SFS). Instructions on how to access the SFS for this course will be sent directly to each student through MavMail approximately 10 days before the end of the term. Each student's feedback via the SFS database is aggregated with that of other students enrolled in the course. Students' anonymity will be protected to the extent that the law allows. UT Arlington's effort to solicit, gather, tabulate, and publish student feedback is required by state law and aggregate results are posted online. Data from SFS is also used for faculty and program evaluations. For more information, visit <http://www.uta.edu/sfs>.

**Final Review Week:** for semester-long courses, a period of five class days prior to the first day of final examinations in the long sessions shall be designated as Final Review Week. The purpose of this week is to allow students sufficient time to prepare for final examinations. During this week, there shall be no scheduled activities such as required field trips or performances; and no instructor shall assign any themes, research problems or exercises of similar scope that have a completion date during or following this week *unless specified in the class syllabus*. During Final Review Week, an instructor shall not give any examinations constituting 10% or more of the final grade, except makeup tests and laboratory examinations. In addition, no instructor shall give any portion of the final examination during Final Review Week. During this week, classes are held as scheduled. In addition, instructors are not required to limit content to topics that have been previously covered; they may introduce new concepts as appropriate.

**Emergency Exit Procedures:** Should we experience an emergency event that requires us to vacate the building, students should exit the room and move toward the nearest exit. When exiting the building during an emergency, one should never take an elevator but should use the stairwells. Faculty members and instructional staff will assist students in selecting the safest route for evacuation and will make arrangements to assist individuals with disabilities.

**Student Support Services:** UT Arlington provides a variety of resources and programs designed to help students develop academic skills, deal with personal situations, and better understand concepts and information related to their courses. Resources include [tutoring](#), [major-based learning centers](#), developmental education, [advising and mentoring](#), personal counseling, and [federally funded programs](#). For individualized referrals, students may visit the reception desk at University College (Ransom Hall), call the Maverick Resource Hotline at 817-272-6107, send a message to [resources@uta.edu](mailto:resources@uta.edu), or view the information at <http://www.uta.edu/studentssuccess/success-programs/programs/resource-hotline.php>

**The IDEAS Center** (2<sup>nd</sup> Floor of Central Library) offers **FREE** [tutoring](#) to all students with a focus on transfer students, sophomores, veterans and others undergoing a transition to UT Arlington. Students can drop in, or check the schedule of available peer tutors at [www.uta.edu/IDEAS](http://www.uta.edu/IDEAS), or call (817) 272-6593.

## Course Schedule

[Topics or the schedule are subject to change]

Week 1.	Course overview
Week 2.	Review of statistics
Week 3.	Similarity Measures
Week 4.	Text Data Modeling
Week 5.	Supervised vs. Unsupervised Models
Week 6.	Instance-based Learning
Week 7.	Decision tree
Week 8.	Evaluation of Classifiers
Week 9.	PCA
Week 10.	Support Vector Machine
Week 11.	Discriminative vs. Generative Models
Week 12.	ANN
Week 13.	ANN
Week 14.	Guest Lectures
Week 15.	Make-up Lectures
Week 16.	Review

---