CIS 4303 – 110, Data Mining

Fall 2023

Texas A&M University-Central Texas

COURSE DATES, MODALITY, AND LOCATION

Course meeting dates: August 28, 2023 – Dec 15, 2023
Class meeting building and room number: Founders Hall 307
Class meeting day and time: Mon 12:30 PM to 1:45 PM, Wed 12:30 PM to 1:45 PM

INSTRUCTOR AND CONTACT INFORMATION

Instructor: Rahul Dwivedi
Office: Founders Hall 323H
Phone: NA
Email: rahul.dwivedi@tamuct.edu (preferred) or Canvas inbox.

Office Hours

Monday and Wednesday: 11 am to 12 noon
Thursday: 12 pm to 1 pm

Students can also email their concerns or questions about any aspect of the course any day of the week. If something needs to be addressed via virtual meeting, students can also request WebEx (https://tamuct.onecampus.com/task/all/webex) or Microsoft Teams meeting(s) during the weekdays.

Student-instructor interaction

I check my university email many times daily and typically respond within 24 hours of receiving weekly emails. Replies during weekends may be delayed up to 48 hours. Email is also the best mode of communication.

This course extensively uses the TAMUCT Canvas Learning Management System (https://tamuct.instructure.com). The course syllabus, schedule, supplemental readings, class announcements, PowerPoint slides, learning modules, homework assignments, exams, and other course-related documents will be posted on Canvas throughout the week. Each student is responsible for the posted material and should check Canvas several times a week for updates.

Emergency Warning System for Texas A&M University-Central Texas

SAFEZONE. SafeZone provides a public safety application that gives you the ability to call for help with the push of a button. It also provides Texas A&M University-Central Texas the ability

to communicate emergency information quickly via push notifications, email, and text messages. All students automatically receive email and text messages via their myCT accounts.

Downloading SafeZone allows access to push notifications and enables you to connect directly for help through the app.

You can download SafeZone from the app store and use your myCT credentials to log in. If you would like more information, you can visit the SafeZone website [www.safezoneapp.com].

To register SafeZone on your phone, please follow these 3 easy steps:
1. Download the SafeZone App from your phone store using the link below:
    o   iPhone/iPad: [https://apps.apple.com/app/safezone/id533054756]
    o   Android Phone / Tablet
        [https://play.google.com/store/apps/details?id=com.criticalarc.safezoneapp]
2. Launch the app and enter your myCT email address (e.g. {name}@tamuct.edu)
3. Complete your profile and accept the terms of service

For updates on COVID information, please monitor the University website [https://www.tamuct.edu/covid19/]

COURSE INFORMATION

Course Overview and description

Examine the theory and application of data mining in accounting, finance, marketing, management, and other business disciplines. Develop basic analytical skills to gain insights and make better decisions. Emphasis on supervised data mining, including k-nearest neighbors, naïve Bayes, and decision trees; unsupervised data mining; and forecasting with time series data.

Course Objective or Goal
Student Learning Outcomes

*   Describe the data mining process, implement similarity measures, assess the predictive performance of data mining models, and conduct principal component analysis.
*   Apply linear probability model and logistic regression model to classify new records.
*   Explain key differences among supervised data mining techniques, apply the k-nearest neighbors and the naïve Bayes method to classify new records.
*   Apply classification trees and ensemble tree models to classify new records and regression trees to predict new records.
*   Conduct hierarchical cluster analysis, k-means cluster analysis, and association rule analysis.
*   Describe the time series forecasting process, use smoothing techniques, linear regression models, nonlinear regression models, and advanced smoothing methods to make forecasts.
*   Apply cross-validation techniques for model selection.

Competency Goals Statements (certification or standards)

None

Required Reading and Textbook(s)

Business Analytics: Communicating with Numbers, Second Edition (loose-leaf or bound edition), McGraw Hill, ISBN for loose-leaf edition: 978-1-265-89710-9. ISBN for bound edition: 978-1-264-30280-2.

It's also acceptable for students who have already taken CIS 3302 – Business Analytics and have a previous edition of the above textbook (but we will not use Connect in this course).

Optional (Reference) Textbook:

R for Data Science, O'Reilly, ISBN: 9781491910399. This text is also available freely as a website: https://r4ds.had.co.nz/

Note: A student of this institution is not under any obligation to purchase a textbook from a university-affiliated bookstore. The same textbook may also be available from an independent retailer, including an online retailer.

Reading Assignments: All assigned chapters and readings will be used for class and canvas discussions. Students are expected to study the assigned readings before each class session.

Supplementary Material: The course textbook will be supplemented with other materials that may include research paper readings, handouts, oral presentations, industry articles, videos including TED talks, research paper readings, case studies, PowerPoint presentations, etc. Students must know how to search for and download research papers/articles from the Web (more specifically from Google Scholar (https://scholar.google.com/)

COURSE REQUIREMENTS

Examinations: There will be three exams. Each exam will be worth 20 points. The exams will have two parts – a multiple-choice part and a programming part. Except for the second exam, the other two will be take-home exams. The multiple-choice portion of the second exam will be an in-class exam (answered online via Canvas), and take-home programming problems are due by the end of the day when it is due (executable code submitted via Canvas).

The exams will all be open book/notes and available via Canvas. You may also use your submitted homework for the in-class part of the exam, but Internet use is not allowed. For the exams, you will be required to know the material from the textbook chapters and material from the class lectures, such as PowerPoint slides and any supplementary/additional material provided. You must be physically present in class to take the multiple-choice portion of the second exam. You must also know how to independently write and debug R programming code for the take-home programming part of the exam.

Individual homework: There will be four individual homework assignments. Students will get approximately two weeks to work on the individual homework assignments. Each homework will be worth 5 points and may cover one or more of the following topics (in no order):

- Basic R programming
- Logistic Regression
- K-nearest neighbor algorithm implementation
- Classification via Naïve Bayes
- Classification and regression trees
- Ensemble tree models
- Association rules and cluster analysis (hierarchical and k-means)

Homework turned in after the due date is considered late. **One point will be deducted for each day the homework is late.** Exceptional circumstances need to be discussed with the instructor ahead of time when possible.

Semester-wide individual research project: One semester-long data mining project will be worth 20 points. Students must choose a freely available secondary data source from the Web (more on this during the second or third week of the class) and propose some research questions. Many secondary data sources are available on the Web; you are free to explore them and are not restricted to a specific type of data or application domain.

The project aims to use the data mining techniques learned during the class to test the verifiability of the proposed research question(s). Students are not required to turn in the data or the programming code but must submit a written report of their findings in a short research paper (guidelines for the team project write-up will be provided soon). Although developing new research problem(s) is not required, students are encouraged to do so.

The individual project is due by the end of the semester, i.e., December 15th, 2023. Exceptional circumstances for any extensions need to be discussed with the instructor ahead of time when possible.

Grading Criteria Rubric and Conversion

Item	Quantity	Points	Total / Percent
Exam(s)	**3**	**20**	**60**
Homework's	**4**	**5**	**20**
Semester-long individual project	**1**	**20**	**20**
Total			**100**

Exams, quizzes, assignments, and an individual term project will receive a numeric score (0-100) each. These scores will be converted to points and totaled to ultimately be converted to letter grades of A, B, C, D, or F, as shown in the example below:

	Percent earned by Student	Max points	Points to be added to the final grade
Exam 1	78%	20	15.6
Exam 2	85%	20	17
Exam 3	90%	20	18
Homework 1	100%	5	10
Homework 2	75%	5	3.75
Homework 3	80%	5	4
Homework 4	90%	5	9
Semester-long individual project	60%	20	12
TOTAL		100	89.35

89.5 – 100 = A
79.5 – 89.49999 = B
69.5 – 79.49999 = C
59.5 – 69.49999 = D
Below 59.5 = F

Grades will not be curved (in the above example, the student with 89.35 will receive a B, NOT an A).

Posting of Grades
All student grades will be posted on the Canvas Grade Book, and students should monitor their grading status through this tool. Grades will be posted within two weeks after the due date.

COURSE OUTLINE AND CALENDAR

Complete Course Calendar

Week	Dates	Chapters	Due at midnight	Notes
1	August 28 – September 1	Faculty Introduction, Syllabus, and R Overview		
2	September 4 – September 8	Chapter 11		Homework 1 available
3	September 11 – September 15	Chapter 11		Individual Project available
September 13: Deadline to drop 16-week classes with no record (Census)				
4	September 18 – September 22	Chapter 9	Homework 1 due	
6	September 25 – September 29	Chapter 9		Homework 2 available
7	October 2 – October 6	Take-home exam one based	Exam 1 due	No class (I am traveling for a

		on chapters 11 and 9		conference).
8	October 9 – October 13	Chapter 12	Homework 2 due	
10	October 16 – October 20	Chapter 12		
12	October 23 – October 27	Chapter 13		Homework 3 available
13	October 30 – November 3	Chapter 13		
14	November 6 – November 10	Chapter 14	Homework 3 due	
November 9: Deadline to drop 16-week classes with a quit (Q) or withdraw (W)				
November 10: Veterans Day				
15	November 13 – November 17	Chapter 14 and Exam 2* based on Chapters 12 and 13		Homework 4 available
16	November 20 – November 24	Chapter 10		
November 23-24: Thanksgiving break				
17	November 27 – December 1	Chapter 10	Homework 4 due	
18	December 4 – December 8	Exam 3 based on chapters 14 and 10		No class.
19	December 11 – December 15		Individual Project Due	No class.

*The multiple-choice portion of the second exam must be taken in class (face-to-face).

Important University Dates
For important university dates from the current Academic Calendar, refer to the link:
https://www.tamuct.edu/registrar/academic-calendar.html

TECHNOLOGY REQUIREMENTS AND SUPPORT

The latest version of R programming language and RStudio integrated development environment needs to be downloaded and installed to successfully complete homework, individual projects, and exams (if you bring in your laptop for class work and taking exams). R and RStudio are open source, freely available from https://www.r-project.org/ and https://www.rstudio.com/products/rstudio/download/. The University's computer labs are equipped with this software.

Technology Requirements

This course will use the A&M-Central Texas Instructure Canvas learning management system. We strongly recommend the latest Chrome, Firefox, Edge, or Safari browsers versions. With a modern web browser, Canvas will run on Windows, Mac, Linus, iOS, Android, or any other device. **Canvas no longer supports any version of Internet Explorer.**

Logon to A&M-Central Texas Canvas [https://tamuct.instructure.com/] or access Canvas through the TAMUCT Online link in myCT [https://tamuct.onecampus.com/] by clicking on the "TAMUCT Online Canvas" tile. You will then log in through our Microsoft portal.

   Username: Your MyCT email address. Password: Your MyCT password

Canvas Support

Use the Canvas Help tab, located at the bottom of the left-hand menu, for issues with Canvas. You can search the support articles or use the Email, Call, or Chat buttons at the bottom of the support pop-up to contact the Canvas Help Desk.

For issues related to course content and requirements, contact your instructor.

Online Proctored Testing

Texas A&M University-Central Texas uses Proctorio for online identity verification and proctored testing. This service is provided at no direct cost to students. If the course requires identity verification or proctored testing, the technology requirements are: Any computer meeting the minimum computing requirements, plus web camera, speaker, and microphone (or headset). Proctorio requires use of the Chrome web browser with their custom plug in installed.

Other Technology Support

For log-in problems, students should contact Help Desk Central, 24 hours a day, 7 days a week.

   Email: helpdesk@tamu.edu
   Phone: (254) 519-5466
   Web Chat: [http://hdc.tamu.edu]
   Please let the support technician know you are an A&M-Central Texas student.

UNIVERSITY RESOURCES, PROCEDURES, AND GUIDELINES

Warrior Center for Student Success
The Warrior Center for Student Success is a comprehensive academic support department at A&M-Central Texas, dedicated to fostering an environment of excellence and empowerment among its student body. The center offers a wide range of programs and services to ensure every student reaches their full potential, and is a haven for students seeking guidance, resources, and a strong support network to excel in their educational journey.

ADA Access and Accommodations: Texas A&M University-Central Texas ensures that students with disabilities have equal access to educational opportunities by providing appropriate

accommodations and support services. If you believe you have a physical, learning or socio-emotional disability requiring reasonable accommodations, please visit Access and Inclusion [https://www.tamuct.edu/student-affairs/access-inclusion.html] for more details or contact the Office of Access and Inclusion, WH-212; (254) 501-5836. Any information you provide is private and confidential.

Success Coaching and Peer Mentoring: Our experienced Success Coaches work one-on-one with students to develop personalized action plans, set academic goals, and build effective study strategies, time management skills, and resilience. Our Peer Mentors provide a valuable support system, offering guidance, encouragement, and a relatable perspective to help students navigate their academic and personal challenges. For more details call 254-501-5836 or 254-501-5928 or visit Academic Support [https://www.tamuct.edu/student-affairs/academic-support.html]. Click the link to schedule a session (virtual or in-person) with a success coach bit.ly/3q7uB50 or visit WH, 111.

Testing Services: We offer a secure and comfortable environment for students and members of the community to take courses and distance learning exams, as well as placement tests and professional certification exams. Our Testing Service also offers resources and support referrals for testing related challenges (test anxiety, learning disabilities, etc.) and supports all approved ADA accommodations. Call (254) 519-5830 or visit the Testing Center [https://www.tamuct.edu/testing-center/].

Tutoring and Supplemental Instruction Services: Our team of qualified Tutors and Supplemental Instructors assist students in various non-writing subjects, promoting academic comprehension and enhancing learning outcomes. Click the link to schedule a tutoring session with a TAMUCT tutor (virtual or in-person) or view tutor availability bit.ly/43Q6wNz. You may also chat live with a remote tutor 24/7 for a variety of subjects through our partnership with Tutor.com, an online tutoring platform that is free to all TAMUCT students. To learn more please visit Tutoring Services [https://www.tamuct.edu/student-affairs/academic-support.html#tutoring] or call (254) 501-5836 or visit the Tutoring Hub in Warrior Hall, 111.

Academic Integrity

Texas A&M University-Central Texas values the integrity of the academic enterprise and strives for the highest standards of academic conduct. A&M-Central Texas expects its students, faculty, and staff to support the adherence to high standards of personal and scholarly conduct to preserve the honor and integrity of the creative community. Any deviation by students from this expectation may result in a failing grade for the assignment and potentially a failing grade for the course. All academic misconduct concerns will be referred to the Student Conduct Office. When in doubt on collaboration, citation, or any issue, please contact your instructor before taking a course of action.

For more information regarding the student conduct process, [https://www.tamuct.edu/student-affairs/student-conduct.html].

If you know of potential honor violations by other students, you may [submit a referral](https://cm.maxient.com/reporting.php?TAMUCentralTexas), [https://cm.maxient.com/reporting.php?TAMUCentralTexas].

Drop Policy

If you discover that you need to drop this class, you must complete the [Drop Request](#) Dynamic Form through Warrior Web.

[https://federation.ngwebsolutions.com/sp/startSSO.ping?PartnerIdpId=https://eis-prod.ec.tamuct.edu:443/samlsso&SpSessionAuthnAdapterId=tamuctDF&TargetResource=https%3a%2f%2fdynamicforms.ngwebsolutions.com%2fSubmit%2fStart%2f53b8369e-0502-4f36-be43-f02a4202f612].

Faculty cannot drop students; this is always the responsibility of the student. The Records and Admissions Office will provide a deadline on the Academic Calendar for which the form must be completed. Once you submit the completed form to the Records and Admissions Office, you must go into Warrior Web and confirm that you are no longer enrolled. If you still show as enrolled, FOLLOW-UP with the Records and Admissions Office immediately. You are to attend class until the procedure is complete to avoid penalty for absence. Should you miss the drop deadline or fail to follow the procedure, you will receive an F in the course, which may affect your financial aid and/or VA educational benefits.

Pregnant and/or Parenting Students Rights and Accommodations

Texas A&M University-Central Texas supports students who are pregnant, experiencing pregnancy-related conditions, and/or parenting. In accordance with requirements of [Title IX and related guidance from US Department of Education's Office of Civil Rights](#), the Associate Dean in the Division of Student Affairs, (254) 501-5909, can assist students who are pregnant, experiencing pregnancy-related conditions, and/or parenting by provide flexible and individualized reasonable accommodations. Students should seek out assistance as early in the pregnancy as possible through the [Pregnancy & Parenting webpage](https://www.tamuct.edu/student-affairs/pregnant-and-parenting-students.html) [https://www.tamuct.edu/student-affairs/pregnant-and-parenting-students.html]. For more information, please visit [Student Affairs](https://www.tamuct.edu/student-affairs/pregnant-and-parenting-students.html) [https://www.tamuct.edu/student-affairs/pregnant-and-parenting-students.html]. If you would like to read more about these [requirements and guidelines](#) online, please visit the website [http://www2.ed.gov/about/offices/list/ocr/docs/pregnancy.pdf].

[Title IX of the Education Amendments Act of 1972](#) specifically prohibits discrimination against a student based on pregnancy, childbirth, false pregnancy, termination of pregnancy, or recovery from any of these conditions [https://www2.ed.gov/about/offices/list/ocr/docs/pregnancy.html].

Students experiencing any form of discrimination due to any of these conditions are encouraged to reach out to the Title IX Coordinator, 254.519.5716, titleix@tamuct.edu, Founders Hall 317B, or the Associate Dean of Student Affairs, 254.501.5909, Warrior Hall 105.

Title IX Rights and Reporting Responsibilities

Texas A&M University-Central Texas is committed to creating a safe and open learning environment for all students. If you or another student has experienced any form of gender discrimination or sexual misconduct, including sexual harassment, sexual assault, dating/domestic violence, and/or sex-based stalking, help and support are available. Our university strongly encourages all members of our campus community to report incidents and seek support for gender discrimination and sexual misconduct through the Title IX Office. You may contact the Title IX Office at 254.519.5716, titleix@tamuct.edu, Founders Hall 317B, or learn more by visiting the Title IX webpage [https://www.tamuct.edu/compliance/titleix.html].

Please be aware that that under Title IX, Texas Senate Bill 212, and System Regulation 08.01.01, [https://policies.tamus.edu/08-01-01.pdf] all university employees are mandated reporters and are required to disclose information about suspected or alleged violations as listed above and defined in System Regulation 08.01.01. If the Title IX Office receives information about an incident, they will reach out to offer information about resources, rights, and procedural options as a member of the campus community. Although I have an obligation to report, you will, in most cases, control how your case will be handled. When working with the Title IX Office you will have access to resources and accommodations but also have the opportunity to express if you wish to move forward with an investigation. Our goal is to make sure you are aware of the options available to you as a student. Community members are not required to respond to this outreach.

If you or another student wishes to speak to a confidential employee who does not have this reporting responsibility, you can contact the Student Wellness & Counseling Center, [https://www.tamuct.edu/student-affairs/student-counseling.html], 254.501.5955, or swacc@tamuct.edu, located in Warrior Hall Room 207L or the Student Support Advocate, 254.501.5978 or ssa@tamuct.edu, located in founder Hall Room 317D.

University Library & Archives

The University Library & Archives provides many services in support of research across campus and at a distance. We offer over 350 electronic databases containing approximately 1,203,947 eBooks and 134,750 journals, in addition to the 96,879 items in our print collection, which can be mailed to students who live more than 50 miles from campus. Research guides for each subject taught at A&M-Central Texas are available through our website to help students navigate these resources. On campus, the library offers technology including cameras, laptops, microphones, webcams, and digital sound recorders.

Research assistance from a librarian is also available 24 hours a day through our online chat service, and at the reference desk when the library is open. Research sessions can be scheduled for more comprehensive assistance, and may take place virtually through WebEx, Microsoft Teams or in-person at the library. Schedule an appointment here

[https://tamuct.libcal.com/appointments]. Assistance may cover many topics, including how to find articles in peer-reviewed journals, how to cite resources, and how to piece together research for written assignments.

Our 27,000-square-foot facility on the A&M-Central Texas main campus includes student lounges, private study rooms, group workspaces, computer labs, family areas suitable for all ages, and many other features. Services such as interlibrary loan, TexShare, binding, and laminating are available. The library frequently offers workshops, tours, readings, and other events. For more information, please visit our [Library website](https://tamuct.libguides.com/index)

[https://tamuct.libguides.com/index]

University Writing Center

Located in Warrior Hall 416 and online, the University Writing Center (UWC) at Texas A&M University–Central Texas (A&M–Central Texas) is a free service open to all A&M–Central Texas students. The face-to-face hours of operation are from 10:00 a.m.-5:00 p.m. Monday and Thursday in Warrior Hall 416. Online tutoring is available Monday thru Thursday from 10:00 a.m.-5:00 p.m. and from 6:00-9:00 p.m. and on Saturdays from 12:00-3:00 p.m.

Tutors are prepared to help writers of all levels and abilities at any stage of the writing process. While tutors will not write, edit, or grade papers, they will assist students in developing more effective composing practices. By providing a practice audience for students' ideas and writing, our tutors highlight the ways in which they read and interpret students' texts, offering guidance and support throughout the various stages of the writing process. In addition, students may work independently in the UWC by checking out a laptop that runs the Microsoft Office suite and connects to WIFI, or by consulting our resources on writing, including all of the relevant style guides. Whether you need help brainstorming ideas, organizing an essay, proofreading, understanding proper citation practices, or just want a quiet place to work, the UWC is here to help!

Students may arrange a one-to-one session with a trained and experienced writing tutor by making an appointment via [WCOnline](https://tamuct.mywconline.com/) [https://tamuct.mywconline.com/]. In addition, you can email Dr. Bruce Bowles Jr. at bruce.bowles@tamuct.edu if you have any questions about the UWC, need any assistance with scheduling, or would like to schedule a recurring appointment with your favorite tutor.

OTHER POLICY STATEMENTS

A Note about Sexual Violence at A&M-Central Texas

Sexual violence is a serious safety, social justice, and public health issue. The university offers support for anyone struggling with these issues. University faculty are mandated reporters, so if someone discloses that they were sexually assaulted (or a victim of Domestic/Dating Violence or Stalking) while a student at TAMUCT, faculty members are required to inform the Title IX Office. If you want to discuss any of these issues confidentially, you can do so through Student Wellness and Counseling (254-501-5955) located on the second floor of Warrior Hall (207L).

Sexual violence can occur on our campus because predators often feel emboldened, and victims often feel silenced or shamed. It is incumbent on ALL of us to find ways to actively create environments that tell predators we don't agree with their behaviors and tell survivors we will support them. Your actions matter. Don't be a bystander; be an agent of change. For

additional information on campus policy and resources visit the Title IX webpage [https://www.tamuct.edu/compliance/titleix.html].

Behavioral Intervention

Texas A&M University-Central Texas cares about the safety, health, and well-being of its students, faculty, staff, and community. If you are aware of individuals for whom you have a concern, please make a referral to the Behavioral Intervention Team. Referring your concern shows you care. You can complete the referral online

[https://cm.maxient.com/reporting.php?TAMUCentralTexas].

Anonymous referrals are accepted. Please see the Behavioral Intervention Team website for more information [https://www.tamuct.edu/bit]. If a person's behavior poses an imminent threat to you or another, contact 911 or A&M-Central Texas University Police at 254-501-5805.

OTHER POLICIES

INSTRUCTOR POLICIES

Policies related to absence, grading, etc.

- Homework, exams, and an individual project must be submitted through Canvas and due on the mid night (11:59 PM) of specified due date/time. No email submissions will be accepted except under unforeseen circumstances.
- Late penalties will be applied to homework submitted after due dates. Refer to late submission policy for individual homework assignments. Late submits are not allowed for exams and an individual project.
- Multiple attempts will not be allowed under any circumstances for any of the exam(s).
- Plagiarism check will be carried out using plagiarism check software built into Canvas available from university. Homework assignments or projects with more than 30% plagiarized material will receive a grade of zero.

My personal statement

- You will receive feedback in the form of graded assignments within two weeks after the due date. I want you to read the feedback that I provide to you (your personal grading notes and Canvas emails).
- I am almost always available via email and typically respond within 24 hours except on weekends. Students may also request virtual meeting(s) during weekdays, if required.
- I prefer email to phone conversations.
- I reserve the right to modify the course syllabus during the semester for the benefit of the students.
- I reserve the right to supplement materials presented in the text with additional course material that may help the students to understand the topic better.

- I reserve the right to modify grading policy rubrics. Any change to grading rubrics will be applied to current and possible future assignments.

Copyright Notice

Students should assume that all course material is copyrighted by the respective author(s). Reproduction of course material is prohibited without consent by the author and/or course instructor. Violation of copyright is against the law and Texas A&M University-Central Texas' Code of Academic Honesty. All alleged violations will be reported to the Office of Student Conduct.