

DSCI 5240/Data Mining and Machine Learning for Business

Instructor Contact

Name: Javier Rubio-Herrero, Ph.D.

Office Location: 357 E - BLB

Phone Number: (940) 565 - 3117

Office Hours: Wednesday 4:30 pm – 6:00 pm, and Thursday 10:45 am – 12:15 pm, to be held via [Zoom](#).

E-mail: javier.rubioherrero@unt.edu

Communication Expectations: The instructor will communicate with students via, in-class announcements, e-mails, Canvas, and office hours. The students will communicate with the instructor via e-mails and office hours.

It is the instructor's aim to answer e-mails within 48 hours of their receipt during business days. Please, be courteous and professional when communicating with your instructor and follow the provided [Online Communication Tips](#). The instructor may not answer e-mails that do not abide to these tips.

Tutor/Grader Contact (This section to be updated shortly)

Tutoring Hours: ITDS Lab Schedules (both BCIS and DSCI) are usually posted during the second week of classes. Check the [Lab Schedule](#) for more information. Your tutoring meetings will happen via [Zoom](#) (meeting ID: 922-325-916). The phone-scannable card with QR code below will also help you to log into your tutoring sessions:



Communication Expectations: The tutor/grader should be your first point of contact for tutoring questions. If you have additional questions, or need additional support for your class needs, you are most welcome to communicate with your instructor after class, during office hours, or via e-mail.

Course Description

This course deals with the problem of extracting information from large databases and designing data-based decision support systems. The extracted knowledge is subsequently used to support human

decision-making in the areas of summarization, prediction, and the explanation of observed phenomena (e.g. patterns, trends, and customer behavior).

Techniques such as visualization, statistical analysis, decision trees, and neural networks can be used to discover relationships and patterns that shed light on business problems. This course will examine methods for transforming massive amounts of data into new and useful information, uncovering factors that affect purchasing patterns, and identifying potential profitable investments and opportunities.

Course Structure

This course will be delivered in remote format and will take place between the weeks of January 11th, 2021 and April 19th, 2020*.

Note: This course employs lecture capture technology to record class sessions. Students may occasionally appear on video. The lecture recordings will be available to you for study purposes and may also be reused in future course offerings.

<i>Week</i>	<i>Topics/Cases</i>	<i>Notes</i>
Week 1 - Introduction	Introduction to Data Mining	Read Sarma Chapter 1-2 Form teams
Week 2 -Working with Data	Data Preparation Linear Regression Quick SAS Enterprise Miner Demo	Read Sarma Chapter 3
Week 3 -Working with Data (ctd.)	Dimension Reduction	Assignment 1 due
Week 4 – Segmentation	Introduction to SAS Enterprise Miner Cluster Analysis	Project proposal due
Week 5 – Segmentation (ctd.)	Hierarchical Cluster Analysis Cluster Analysis in SAS Enterprise Miner	Assignment 2 due
Week 6	Midterm exam	No class meeting
Week 7 – Classification	Introduction to Classification Naïve Bayes	Assignment 3 due
Week 8 – Classification(ctd.)	Logistic Regression	Read Sarma Chapter 6
Week 9 – Classification (ctd.)	Decision Trees and Random Forests	Read Sarma Chapter 4 Project status report due
Week 10 – Classification (ctd.)	Neural Networks Classification in SAS Enterprise Miner	Read Sarma Chapter 5
Week 11 – Classification (ctd.)	Evaluating Classifiers	Read Sarma Chapter 7
Week 12 – Database Methods	Association Rules Recommender Systems	Assignment 4 due
Week 13 – Text Mining	Text Mining	Read Sarma Chapter 9 Project report due Project presentation due

* Course structure subject to change in order to accommodate specific circumstances that may appear throughout the semester.

<i>Week</i>	<i>Topics/Cases</i>	<i>Notes</i>
Week 14 – Wrap up	Data Mining in Real Life (Video)	No class meeting
Week 15 – Wrap up (ctd.)	Project Presentations	
Week 16 – Wrap up (ctd.)	Final Exam	No class meeting

A list of important dates for this semester can also be consulted at the [Spring 2021 Registration Guide](#).

Course Prerequisites or Other Restrictions

There are no pre-requisites for this class. Some knowledge of basic statistics and probability will be helpful.

Course Objectives

1. Students will be able to articulate basic data mining and machine learning concepts and their application in the business context.
2. Students will be able to use analytical software to identify and interpret complex patterns in multidimensional data.
3. Students will be able to describe and use common descriptive and predictive algorithms.
4. Students will be able to describe all phases of decision making, including data discovery and capture, data analysis and confirmation, presentation of results, and implementation of results.

Technical Requirements & Skills

All students are responsible satisfying the requirements below and they must make sure that they comply with them when they perform the tasks that are needed for the successful completion of this course.

- Computer with Webcam
- Reliable internet access
- Speakers
- Microphone
- Microsoft Office Suite
- [Canvas Technical Requirements](#)
- LockDown Browser and Respondus Monitor
- Business or scientific calculator. **Only these calculators will be allowed in quizzes and exams. Calculators from phones, tablets or other electronic devices will not be allowed.**

Textbooks and Materials

1. **Textbooks:** Kattamuri S. Sarma, *Predictive Modeling with SAS Enterprise Miner: Practical Solutions for Business Applications*, 3rd Edition, SAS Institute 2017, ISBN: 978-1629602646. RECOMMENDED, NOT REQUIRED.
2. **Software:** SAS Enterprise Miner – Available at the CoB lab or via the cloud.
3. **Computer:** Access to a computer with an operational webcam and microphone is required. More details will follow in the section [Technical Requirements & Skills](#).
4. **Calculator:** A simple scientific calculator will be required for exams. No programmable or graphing calculators are permitted. No sharing of calculators during exams will be allowed.

Teaching Philosophy

Since this course will be taught remotely, we will aim at resembling a face-to-face class environment as much as possible. To make the class more dynamic, I will opt for combining slides and hand-written notes. I will write notes on-the-go in class as I explain concepts and exercises. This course has a clear mathematical background and mathematics require work and patience to be understood. I believe that this understanding cannot be attained without writing things down by yourselves and, for this reason, I expect you to be active in taking notes and, if needed, revisit the recorded lectures.

Grading

Grades measure the performance of a student in individual courses. Students will be evaluated based on their performance and command of the course materials. Final grades are broken down as follows[†]:

Course Requirement	Points	Percentage of Final Grade
Participation/ Professionalism	50	5%
Homework assignments	200	20%
Exam 1	250	25%
Exam 2	250	25%
Group Project [‡]	250	25%
Total	1000	100%

A student's final grade will be determined following the grading scale[§] below:

A = 900 – 1000 points

B = 800 – 899 points

[†] The breakdown below is contingent upon acceptance of this syllabus and its addenda. Each student submission will receive a 0 until such acceptance is received by the instructor.

[‡] PhD students should form groups together and are encouraged to select a project that could result in a conference proceeding or a journal publication.

[§] This grading scale is tentative and may be adjusted at the end of the semester based on overall class performance.

C = 700 – 799 points

D = 600 – 699 points

F = 0 – 599 points

Course Policies

Attendance Policy and Class Participation

I expect, but do not enforce, that students attend each class session. Final grades in the course are determined, in part, by participation in course discussions and assignments. Regardless of attendance, **students are responsible for learning the contents, meet the deadlines, and follow the directions explained in class.**

COVID-19 Impact on Attendance

While attendance is expected as outlined above, it is important for all of us to be mindful of the health and safety of everyone in our community, especially given concerns about COVID-19. Please contact me if you are unable to attend class because you are ill, or unable to attend class due to a related issue regarding COVID-19. It is important that you communicate with me prior to being absent so I may make a decision about accommodating your request to be excused from class.

If you are experiencing any [symptoms of COVID-19](https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html) (https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html) please seek medical attention from the Student Health and Wellness Center (940-565-2333 or askSHWC@unt.edu) or your health care provider PRIOR to coming to campus. UNT also requires you to contact the UNT COVID Hotline at 844-366-5892 or COVID@unt.edu for guidance on actions to take due to symptoms, pending or positive test results, or potential exposure. While attendance is an important part of succeeding in this class, your own health, and those of others in the community, is more important.

Announcements Policy

In addition to in class announcements, important messages may be communicated via [Canvas](http://canvas.unt.edu) (http://canvas.unt.edu). You are responsible for actively monitoring Canvas for announcements which, among other things, may be related to changes in schedule, syllabus, or other key aspects of the course.

Group Participation Policy

This course requires considerable group work. Details related to team formation will be delivered in class and/or on the course website. Because of the volume of group work and the possibility that some team members may not contribute to the overall team success, a team may document problematic behaviors and recommend that a non-performing member of the team be “fired.” In such circumstances, I will meet with the team and create a performance improvement plan (PIP) for the offending team member. **If the team member fails to meet the obligations outlined in the PIP, or if the problematic behaviors reoccur after completion of the PIP, they will be removed from the team and must complete all remaining deliverables alone. The maximum grade that may be achieved on a given deliverable by an individual fired from their team is 80% of the associated points.**

Assignment Policy

Homework assignments will be announced via Canvas. The due date for each homework will be announced on the posting date along with extra instructions (if needed). As a rule of thumb, homework has to be completed professionally and therefore it needs to be well presented, clean, readable, and easy to follow. The instructor and the grader may reduce your grade at their discretion if these general guidelines are not correctly followed.

All assignments and project deliverables must be submitted through Canvas. Alternative submission methods (e.g., paper, email, USB drive) are not acceptable unless prior permission of the instructor is obtained. **Any assignment that is submitted after the submission deadline will not be graded and will result in a zero on that assignment.** Consequently, make sure that you do not leave your submissions for the last minute. See the [Late Work](#) subsection for more details about late submissions. **Make sure that the names of all the students that participated in the homework are listed in the first page of your Word file. Students that participated in an assignment that are not listed will not receive any credit for this submission.**

Students are welcome to review both assignments and exams after grading to improve their understanding of course material and check for the presence of grading errors. This review must be conducted within two weeks of the assignment deadline or exam. No changes to grades will be made after this time period.

Examination Policy

The following policies will apply to your exams:

- All exams be closed-book and closed-note.
- You will need a desktop or a laptop to complete your exams. If you use a laptop, make sure you connect it to the power grid. **I will not repeat any exam if your laptop runs out of battery.**
- Phones and tablets will **not** be allowed.
- Students are responsible for their materials on an exam. Loaning or sharing any materials is strictly prohibited.
- It is **strictly forbidden** that students share information during an exam. Failure to comply with this will result in disciplinary actions as described in the [Academic Integrity Policy](#) subsection of this syllabus.
- Students who fail to bring their own required materials will take the exams without them and to the best of their ability.
- Further instructions will be announced ahead of time, if needed.
- In case you lose Internet connection during an exam:
 - If you are having a short disruption, Canvas will allow you to keep on working on the exam and will reconnect automatically once your Internet service resumes.
 - If the issues persist, contact the [Student Helpdesk](#) (helpdesk@unt.edu or 940.565.2324) immediately and document the remedy ticket number.
 - If the Student Helpdesk cannot address your problems or if they advise you to contact your instructor, immediately report this disruption to both your instructor and the grader and enclose a copy of the current state of your exam. This is very important to keep a recorded time stamp of the events.

- If the issues still persist, complete your exam offline and send it via e-mail to both your instructor and the grader. **Important:** always do this within the allotted time for the exam.
- Missed exams will receive zero marks unless a legitimate excuse is presented, as described in the [Late Work](#) subsection of this syllabus.
- In the case that you want to review a midterm exam that you took, this request must be done to the grader/instructor within two weeks since the grades were posted. Later reviews will not be accepted.
- If you want to review a final exam, this request must be done to the grader/instructor within 10 days or before the grades are due at the Registrar's Office, whatever is shorter.

Late Work

It is **highly recommended** that you work on your submissions well in advance to avoid last-minute issues. I will **not** accept any late submissions nor agree to make-up tests or exams except in the following cases:

- Medical emergency cases, in which case a doctor's note is required.
- Family emergency cases, in which case a written verifiable proof is required.
- Any other excused absence specified in Chapter 6 of Faculty Affairs document, [Student Attendance and Authorized Absences](#).

Late submissions and make-up tests or exams must always be agreed **before the corresponding due date**. No makeup tests or exams will be offered due to poor performance in exams.

Syllabus Change Policy

The contents of this syllabus might be changed to the instructor's discretion in order to adjust the course to the specific circumstances of each semester.

Rules of Engagement

Rules of engagement refer to the way students are expected to interact with each other and with their instructors online. Here are some general guidelines:

- Treat your instructor and classmates with respect in any communication online or face-to-face, even when their opinion differs from your own.
- Always use your professors' proper title: Dr. or Prof., or if in doubt use Mr. or Ms.
- Unless specifically invited, don't refer to your instructor by first name.
- Use clear and concise language.
- Remember that all college level communication should have correct spelling and grammar (this includes discussion boards).
- Avoid using "text-talk" unless explicitly permitted by your instructor.
- Avoid using all caps while communicating digitally. This may be interpreted as "YELLING!"
- Be careful with personal information (both yours and other's) and avoid sending confidential information via e-mail.
- Be cautious when using humor or sarcasm in emails or discussion posts as tone can be difficult to interpret digitally.
- Keep in mind that online posts can be permanent, so think first before you type.

See these [Engagement Guidelines](https://clear.unt.edu/online-communication-tips) (https://clear.unt.edu/online-communication-tips) for more information.

Course Evaluation

The Student Perceptions of Teaching (SPOT) is the student evaluation system for UNT and allows students the ability to confidentially provide constructive feedback to their instructor and department to improve the quality of student experiences in the course. At some stage during this course, students will be given access to this evaluation. I would certainly appreciate your constructive feedback when the time comes.

Getting Help

Technical Assistance

Part of working in the online environment involves dealing with the inconveniences and frustration that can arise when technology breaks down or does not perform as expected. Here at UNT we have a Student Help Desk that you can contact for help with Canvas or other technology issues.

UIT Help Desk: [UIT Student Help Desk site](http://www.unt.edu/helpdesk/index.htm) (http://www.unt.edu/helpdesk/index.htm)

Email: helpdesk@unt.edu

Phone: 940-565-2324

In Person: Sage Hall, Room 130

Walk-In Availability: 8am-9pm

Telephone Availability:

- Sunday: noon-midnight
- Monday-Thursday: 8am-midnight
- Friday: 8am-8pm
- Saturday: 9am-5pm

Laptop Checkout: 8am-7pm

For additional support, visit [Canvas Technical Help](https://community.canvaslms.com/docs/DOC-10554-4212710328) (https://community.canvaslms.com/docs/DOC-10554-4212710328)

Computer Labs

The BLB computer labs on the first floor will open with reduced hours and reduced seating to comply with social distancing guidelines. They will not be BLB laptop checkouts due to potential contamination issues. The library [has computers to check out for up to 24 hours](#) and the [RCOB virtual computer labs](#) will be open for business continuously.

UNT Policies

Academic Integrity Policy

Academic Integrity Standards and Consequences. According to [UNT Policy 06.003, Student Academic Integrity](#), academic dishonesty occurs when students engage in behaviors including, but not limited to cheating, fabrication, facilitating academic dishonesty, forgery, plagiarism, and sabotage. A finding of

academic dishonesty may result in a range of academic penalties or sanctions ranging from admonition to expulsion from the University.

Academic dishonesty is an extremely serious issue. It will not be tolerated and will be prosecuted according to UNT Policy 06.003. You are responsible for knowing what those behaviors above (cheating, plagiarism, etc.) mean and when you might be incurring any of them.

ADA Policy

UNT makes reasonable academic accommodation for students with disabilities. Students seeking accommodation must first register with the Office of Disability Accommodation (ODA) to verify their eligibility. If a disability is verified, the ODA will provide a student with an accommodation letter to be delivered to faculty to begin a private discussion regarding one's specific course needs. Students may request accommodations at any time, however, ODA notices of accommodation should be provided as early as possible in the semester to avoid any delay in implementation. Note that students must obtain a new letter of accommodation for every semester and must meet with each faculty member prior to implementation in each class. For additional information see the [ODA website \(https://disability.unt.edu/\)](https://disability.unt.edu/).

Emergency Notification & Procedures

UNT uses a system called Eagle Alert to quickly notify students with critical information in the event of an emergency (i.e., severe weather, campus closing, and health and public safety emergencies like chemical spills, fires, or violence). In the event of a university closure, please refer to Canvas for contingency plans for covering course materials.

Retention of Student Records

Student records pertaining to this course are maintained in a secure location by the instructor of record. All records such as exams, answer sheets (with keys), and written papers submitted during the duration of the course are kept for at least one calendar year after course completion. Course work completed via the Canvas online system, including grading information and comments, is also stored in a safe electronic environment for one year. Students have the right to view their individual record; however, information about student's records will not be divulged to other individuals without proper written consent. Students are encouraged to review the Public Information Policy and the Family Educational Rights and Privacy Act (FERPA) laws and the University's policy. See UNT Policy 10.10, Records Management and Retention for additional information.

Acceptable Student Behavior

Student behavior that interferes with an instructor's ability to conduct a class or other students' opportunity to learn is unacceptable and disruptive and will not be tolerated in any instructional forum at UNT. Students engaging in unacceptable behavior will be directed to leave the classroom and the instructor may refer the student to the Dean of Students to consider whether the student's conduct violated the Code of Student Conduct. The University's expectations for student conduct apply to all instructional forums, including University and electronic classroom, labs, discussion groups, field trips, etc. Visit UNT's [Code of Student Conduct \(https://deanofstudents.unt.edu/conduct\)](https://deanofstudents.unt.edu/conduct) to learn more.

Access to Information - Eagle Connect

Students' access point for business and academic services at UNT is located at: my.unt.edu. All official communication from the University will be delivered to a student's Eagle Connect account. For more

information, please visit the website that explains Eagle Connect and how to forward e-mail [Eagle Connect](https://it.unt.edu/eagleconnect) (<https://it.unt.edu/eagleconnect>).

Student Evaluation Administration Dates

Student feedback is important and an essential part of participation in this course. The student evaluation of instruction is a requirement for all organized classes at UNT. The survey will be made available during weeks 13, 14 and 15 of the long semesters to provide students with an opportunity to evaluate how this course is taught. Students will receive an email from "UNT SPOT Course Evaluations via IASystem Notification" (no-reply@iasystem.org) with the survey link. Students should look for the email in their UNT email inbox. Simply click on the link and complete the survey. Once students complete the survey they will receive a confirmation email that the survey has been submitted. For additional information, please visit the [SPOT website](http://spot.unt.edu/) (<http://spot.unt.edu/>) or email spot@unt.edu.

Class Recordings & Student Likenesses

Synchronous (live) sessions in this course will be recorded for students enrolled in this class section to refer to throughout the semester. Class recordings are the intellectual property of the university or instructor and are reserved for use only by students in this class and only for educational purposes. Students may not post or otherwise share the recordings outside the class, or outside the Canvas Learning Management System, in any form. Failing to follow this restriction is a violation of the UNT Code of Student Conduct and could lead to disciplinary action.