CSCE 5215: Machine Learning, University of North Texas

Instructor Contact

Name: Jing Yuan  
Pronouns: she/her  
Office Location: Discovery Park F276  
Office Hours: Wednesday/Friday 09:00 am - 10:00 am or by appointment  
Email: jing.yuan@unt.edu

TA: Shanmukha Datta Koganti  
Office Hours: Monday 1:00 pm - 2:30 pm, Friday 1:30 pm - 3 pm or by appointment  
Office Location: CSE Help Lab, Cubicle TBD  
Email: shanmukhadattakoganti@my.unt.edu

TA: Ashraf Syed  
Office Hours: Tuesday/Thursday 3:00 pm - 4:30 pm or by appointment  
Office Location: CSE Help Lab, Cubicle TBD  
Email: ashrafsyyed@my.unt.edu

Welcome to UNT!

As members of the UNT community, we have all made a commitment to be part of an institution that respects and values the identities of the students and employees with whom we interact. UNT does not tolerate identity-based discrimination, harassment, and retaliation. UNT’s full Non-Discrimination Policy can be found in the UNT Policies section of the syllabus.

Course Description

Machine learning is the process of applying algorithms to learn directly from data to make predictions and decisions without being explicitly programmed. Topics include a wide variety of supervised learning methods, both regression and classification, with an emphasis on those that perform well on large feature sets. Ensemble methods are used to combine independent approaches efficiently. Unsupervised and semi-supervised methods will demonstrate the power of learning from data without an explicit training target or goal.

Learning outcomes: Students in this course will learn how to apply sophisticated algorithms to large data sets, with a focus on practical application. The goal will be to create models that can make automated predictions or classifications on new data, or make inferences on unlabeled data to aid in understanding and future prediction models.
Communication Expectations

You will be expected to regularly check university email and attend class in-person regularly. When you miss a class, you are expected to check Canvas shortly after class to be aware of assignments, quizzes, and other materials. Questions not answered in class are best asked before or after class. For in-depth assistance on course content, you are expected to meet with the TAs prior to meeting with the instructor. For quick questions, email is preferred, and you can expect a response within 24 hours during the work week (M-F). For involved questions or discussions not appropriate for the TAs, office hours are preferred.

Course Objectives

By the end of the course, students will be able to:

1. Properly collect and organize data to extract relevant features for learning
2. Build predictive models, both regression and classification, using a variety of modeling strategies
3. Use unsupervised learning techniques to understand high dimensional data sets
4. Apply advanced techniques such as reinforcement learning for robust behavior in complex, changing environments

Course Prerequisites

- CSCE 3110: Data Structures and Algorithms with a grade of C or better
- Experience with Python is beneficial as it is used extensively in the course, but significant prior programming experience with any language will be sufficient

Course Structure

**Time:** Friday 10:00am – 12:45pm  
**Location:** DP B185  
**Credit hours:** 3  
**Dates:** August 21 - December 15, 2023

Tentative topics

- Machine Learning concepts
- Supervised learning
- Model validation and selection
- Feature selection and feature engineering
- Ensemble methods: bagging, boosting
- Unsupervised learning: PCA, ICA
- Clustering: K-means and DBSCAN
- Semi-supervised learning
- Reinforcement learning
- Bayesian networks and Markov models
• Deep learning

Materials

Readings will all be fairly dense, so please search for additional resources (e.g. wikipedia, coursera lectures) as needed. All attempts will be made to provide sufficient resources for everyone.

Textbook: Although no text is required, the following textbook is highly recommended:
• Additional online resources (however many other resources are available)
  ○ Elements of Statistical Learning http://statweb.stanford.edu/~tibs/ElemStatLearn/

Course communication: We will be using the Canvas discussion board
• Feel free to use the forum to ask questions of the group, ask about partners for problem sets, or to make comments that the rest of the class might find useful.
• The forum is primarily for timely, supplementary communication.
• Contact the instructor if you are not available to access the discussion forum after the second week of the course.

Technical Requirements and Skills

Minimum Technology Requirements
• Computers are required for exams
  ○ In-class exams: You will need to bring a laptop on exam days. These will be done individually with LockDown Browser on your computer in class. You will be expected to connect to the UNT wireless network.
  ○ Quizzes and exams will use the Canvas quiz system
    ■ Canvas Basics for UNT Students (https://online.unt.edu/canvas-basics-unt-students)
    ■ Canvas Technical Requirements (https://clear.unt.edu/supported-technologies/canvas/requirements)
• Computers are optional during class, outside of exam times
  ○ There will be occasional in-class activities along with lecture that are not required, but may help in understanding and applying the material
• Students will be expected to Download and install Anaconda Python version 3 and be able to open a Jupyter notebook

Technical Assistance
UNT IT Help Desk (http://www.unt.edu/helpdesk/index.htm)
Email: helpdesk@unt.edu
Live Chat: https://aits.unt.edu/support/chat
Phone: 940-565-2324
**In Person:** Sage Hall, Room 330  
**Hours and Availability for Walk-In, Telephone, and Laptop Checkout:** visit [https://aits.unt.edu/support](https://aits.unt.edu/support) for up-to-date hours and availability

For additional support, visit Canvas Technical Help (http://community.canvaslms.com/docs/DOC10554-4212710328)

## Course Requirements

### Readings, Quizzes, and Exams

**Course Philosophy:** In this course you will be evaluated more often than other courses, with weekly quizzes, assignments, and three exams. Consider the points given to each as a guide to the effort expected. Of particular note: quizzes will be “light-weight” and you are allowed to take them at home. Given the amount of material, it is suggested that for assignments and quizzes you focus on being succinct, and for readings you focus on the main issues.

**Readings/Tutorials:** Generally, quizzes will be given based on the readings and lecture material. As will be clear in the first few weeks, quizzes will test your knowledge on the most important aspects of the readings only.

**Canvas take-home quizzes:** These quizzes are meant to focus students on the important aspects of the readings or lectures. You will be allowed to take these quizzes online. All canvas quizzes will be due the last day of class, but it is suggested that you finish them in the suggested period in preparation for discussions and exams.

**Exams:** Exam days are already posted and are considered fixed. Prior arrangements can potentially be made without loss of points, but have to be discussed. Missed exams: Exams cannot be missed without prior arrangements or later documented proof of extenuating circumstances.

### Assignments

Assignments are designed to engage you in your learning, so you can begin to apply these principles in practice and tailor them to your needs.

**Assignments** are generally due at the end of the day one week after they are assigned, unless otherwise specified. Reports and presentation slides are to be turned in as PDF. Code is to be turned in with both Jupyter notebook and PDF form, along with any files necessary to run your assignment. Results should be presentable, with appropriate comments for someone to follow what you have done. Assignments are to be turned in individually, although students are encouraged to work together extensively. Assignments will have minimal impact on grades - they are for learning and self-evaluation rather than grading.
Short presentations: There are far more tools and techniques than we can possibly cover, many which may be particularly relevant to your interests. Also, the ability to distill complex topics into a form useful for the audience is a critical skill to develop. Students will be expected to present one concept, tool, or technique which goes beyond what is covered in the course. Each presentation is to be 5-10 minutes on the whiteboard or with slides. These presentations will be concentrated on 1-2 days in the semester. Only a basic familiarity is expected of students observing the presentations.

Projects: After a few weeks into the course, you will select a collaborative project. Project proposals, progress reports, and final reviews will be part of the process. You are required to work in groups, as this is part of a full and complete education. All people in the group are expected to contribute. This is your opportunity to demonstrate what you have learned in a way that reaches beyond the selection of tools, data sets, and approaches demonstrated in the course. Commonly students find a unique, complex data set and associated learning problem and apply the techniques presented in the class. The goal here is to create a coherent, completed project for presentation at the end of class. Essentially ask yourself what you would want to show an employer (or brag about to others) demonstrating what you have learned in the course.

Grading

The expected distribution of percentage is given below, this is subject to minor modification based on actual points given. Note, due to the nature of the course, exams and quizzes are a significant means of establishing your final grade, so please complete the assignments in a timely way and study appropriately prior to each quiz and exam.

- Assignments and take-home quizzes: 30%
  - Canvas take-home quizzes: 5%
  - Assignments: 25%
- Presentation and Project: 25%
  - Short presentations: 1%
  - Projects: 24%
    - Proposal: 2%
    - Update: 2%
    - Report: 10%
    - Project presentation: 10%
- Exams: 45%
  - Exam I: 10%
  - Exam II: 10%
  - Final Exam: 25%

Grading Scale: A=above 90%, B=80%-89.99%, C=70%-79.99%, D=60%-69.99%, F=below 60%. No exceptions.
Course Evaluation

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. Spot evaluations will be available November 20 - December 7th, 2023

Course Policies

Examination Policy
Exams will be on the computer using the Canvas quiz system. You need to bring a laptop on the appropriate exam days. Exams are “closed book” - no use of materials outside the canvas exam system unless otherwise specified in advance. Quizzes and exams must be taken in the classroom unless special accommodations have been made through the Office of Disability Accommodation (ODA). Another other accommodation must be given by prior arrangement with the instructor, otherwise documentation proving an extenuating circumstance will have to be provided after the missed exam. Time will be limited, and all work will be individual. You are strongly encouraged to attempt to solve the tasks iteratively and incrementally - write code that works first, but works poorly, and improve from there, rather than write perfect code top to bottom. Exams will focus on the most recent material but are expected to be cumulative in scope.

Technical errors during exams
If during an online quiz or exam there is a technical error which affects your ability to complete the assignment, you are immediately to let the quiz or exam proctor know and the instructor will discuss ways to allow you to resume the test without giving an unfair advantage. In the event of any unexpected server outage or any unusual technical difficulty which prevents students from completing a time sensitive assessment activity, the instructor will extend the time windows and provide an appropriate accommodation based on the situation. Students should immediately report any problems to the instructor.

Late Policy
UNT instructors have the prerogative to accept or not to accept late work. All assignments are due by specified date at 11:59 PM Central Time. Note that 10% of the grade will be deducted for each day the assignment is late. You will get 0 grade if the assignment is 2 days (48 hours) late. A request for missing submissions may be sent at the discretion of the TA or instructor, but is not guaranteed, and a reduction of points may also occur in a way that is consistent for the rest of the class. Given the frequency of assignments and quizzes, and this flexible late policy, there may be a significant lag between submission and entry into the grade book for already-documented small-point assignments.

Attendance Policy
You are expected to attend lectures and to complete all readings, however, this course does not use participation points and there is no penalty for missing days without exams, or group project efforts. There is no need to let the instructor know you have missed a class, however, you are responsible for keeping up with the material covered in the class if you are not present. If a
class is missed, you are expected to proactively reach out to classmates, the TA, or the instructor if there are any questions.

Individual attendance is not required except on group project and exam days. There is no direct participation grading, but in the past there has been a strong correlation between engagement and accomplishment in courses - especially for those that are struggling with the material. Feel free to prioritize your time, but prioritize wisely.

**Syllabus Change Policy**

Any substantial changes to the syllabus after the first week will be highlighted in red on the online platform. Approximate point values are expected to vary but will be fixed when the assignment or exam is given.

**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.

**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/).

**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 Blackboard 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 Blackboard 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) 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).

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/) or email spot@unt.edu.

Getting Help

Student Support Services

- Registrar (https://registrar.unt.edu/registration)
- Financial Aid (https://financialaid.unt.edu/)
- Student Legal Services (https://studentaffairs.unt.edu/student-legal-services)
- Career Center (https://studentaffairs.unt.edu/career-center)
- Multicultural Center (https://edo.unt.edu/multicultural-center)
- Counseling and Testing Services (https://studentaffairs.unt.edu/counseling-and-testing-services)
- Student Affairs Care Team (https://studentaffairs.unt.edu/care)
- Student Health and Wellness Center (https://studentaffairs.unt.edu/student-health-and-wellness-center)
- Pride Alliance (https://edo.unt.edu/pridealliance)
Academic Support Services

- Academic Resource Center (https://clear.unt.edu/canvas/student-resources)
- Academic Success Center (https://success.unt.edu/asc)
- UNT Libraries (https://library.unt.edu/)
- Writing Lab (http://writingcenter.unt.edu/)
- MathLab (https://math.unt.edu/mathlab)