TEXAS
The University of Texas at Austin

INF 385T SPECIAL TOPICS IN INFORMATION SCIENCE: FOUNDATIONS OF DATA SCIENCE

27800

Fall 2024

CLASS MEETS:

03:00 PM - 06:00 PM

UTA 1.208

Instructor: Dr. Shounak Roychowdhury

Office: UTA 5.448

Pronouns: he

Email: shounak.roychowdhury@ischoool.utexas.edu

Office hours: 12:30 PM to 1:30 PM, Wednesday or by appointment.

Course Description

UNIVERSITY CATALOG COURSE DESCRIPTION

This class explores various data science models, both traditional and the state-of-the-art techniques. The course is designed to provide mathematical and computational basis such as Linear Algebra, Optimization techniques, and probabilistic modeling for different types of machine learning models. The goal of the class is providing a foundational basis for data science techniques. The class focuses on PSETs and a final data science project.

The proliferation of open-source and proprietary data sciences packages has enabled the rise of low-code and no-code data science practices. This level of abstraction is undoubtedly helpful to users with limited mathematical knowledge. However, the data science packages are primarily computational packages that heavily rely on sophisticated algorithmic techniques based on different branches of mathematics, such as linear algebra, probabilistic theory, and convex optimization. To better understand data science and contribute to its research in theory and practice, it becomes imperative to know and understand these broad mathematical concepts in the context of various tools, techniques, and algorithmic designs. This course discusses and explores different data science techniques and their applications in light of the mathematical frameworks and Python libraries. In addition, the course examines essential survey papers that provide a comprehensive landscape of different topics.

TEXAS
The University of Texas at Austin

PRE-REQUISITES FOR THE COURSE

We will not teach the Python programming. You are expected to know how to code in Python, which you will use in your PSETs.

LEARNING OUTCOMES

1. Get foundational concepts like Linear Algebra, Probability, Optimization that are required for solving data science related problems, a critical precursor to effective collaborations in industry or academia.
2. Identify types of algorithms commonly used to solve each problem alongside their general properties that make them well-suited for the problem.
3. Conduct research in machine learning to enhance their expertise on a topic of their choice. Deliver an oral presentation that explains the research. Review fellow students' presented research and provide constructive feedback. Communicate the research through a final report.

FLAG COURSES

How Will You Learn?

STATEMENT OF LEARNING SUCCESS

Your success in this class is important to me. We all learn differently, and everyone struggles sometimes. You are not, ever, the only one having difficulty! If there are aspects of this course that prevent you from learning or exclude you, please let me know as soon as possible. Together we will develop strategies to meet both your needs and the requirements of the course. I also encourage you to reach out to the student resources available through UT and I am happy to connect you with a person or on Zoom.

TEACHING MODALITY INFORMATION

This is in-person class that requires attendance and encourages participation. If there will be no alternative to in-person attendance, other than normal emergency accommodations, note this clearly. The lectures will not be recorded or posted anywhere. The slides for each class will be uploaded to the File folder of canvas. There are two parts to this course: 1) a set of lectures given by the instructor, and 2) paper and project presentation by students of various topics of computer vision followed by 14-16 pages report. There are no programming assignments or labs.

COMMUNICATION

The course Canvas site can be found at utexas.instructure.com. Please email me through Canvas. You are responsible for ensuring that the primary email address you have recorded with the university is the one you will check for course communications because that is the email address that Canvas uses.

TEXAS
The University of Texas at Austin

ASKING FOR HELP

I will be available for an hour for office hours right after the class. Otherwise, send me an email or talk to me before or after the class.

DIVERSITY, EQUITY AND INCLUSION

It is my intent that students from all diverse backgrounds and perspectives be well served by this course, that students' learning needs be addressed, and that the diversity that students bring to this class can be comfortably expressed and be viewed as a resource, strength and benefit to all students. Please come to me at any time with any concerns.

DISABILITY & ACCESS (D&A)

The university is committed to creating an accessible and inclusive learning environment consistent with university policy and federal and state law. Please let me know if you experience any barriers to learning so I can work with you to ensure you have equal opportunity to participate fully in this course. If you are a student with a disability, or think you may have a disability, and need accommodations please contact Disability & Access (D&A). Please refer to the D&A website for more information: http://diversity.utexas.edu/disability/. If you are already registered with D&A, please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations and needs in this course.

Course Requirements and Grading

REQUIRED MATERIALS

Moritz Hardt and Benjamin Recht, "Patterns, Predictions, and Actions: Foundations of Machine Learning", Princeton University Press, 2023

Kevin Murphy, "Machine learning: A probabilistic perspective", MIT Press, 2013

Hastie, Tibshirani, and Friedman, "The elements of statistical Learning: Data Mining, Inference, and Prediction", Springer Verlag, 2009

REQUIRED DEVICES

CLASSROOM EXPECTATIONS

- **Class attendance** -- This class requires attendance because there are **no video recordings for the lectures**. The lectures are interactive, and participation is highly encouraged.

- **Class participation --** I will take attendance in every class. Attendance amounts to 5% of your grade.

- **Behavior expectations** **No food or drinks in the class**. Please **silence your cell phones** during the class. I do not permit any for recording, either audio or video in the class. Also, see formally and officially, Section 11-400 of the Institutional Rules in the GIC.

- **Professional Standards** Respect your peers and instructor. Keep a well-balanced attitude in the class so that we call benefit learning from each other.

PROGRAMMING

The PSETs require coding in Python. It is expected that the student knows how to code in Python and Python Libraries.

Again, as said before, this class will not teach Python programming.

CLASS STRUCTURE

The following table represents how you will demonstrate your learning and how we will assess the degree to which you have done so.

Assignments	Percent of Total Grade
1. Attendance and Participation	5
3. Three programming PSETs: Python coding	30
4. Midterm	25
5. Research paper discussion and presentation in class	10
6. Final project presentation and report in class	10 (presentation) + 20 (report)

EXAM

There will be one late midterm in the class. It will not be offered at different times to accommodate any personal travel, interviews, rest days, etc.

LATE WORK AND MAKING UP MISSED WORK

If you turn in your assignment late, expect points to be deducted. Extensions will be considered on a case-by-case basis, but in most cases, they will not be granted.

ABSENCES

Inform the instructor if you intend to be absent.

TEXAS
The University of Texas at Austin

USE OF A CURVE

Depends on the overall performance of the class.

EQUITABLE ACCOMMODATION

N/A

EXTRA CREDIT

N/A

+/- GRADNG POLICY

I will use +/- grades will be used for the final class grade.

GRADE BREAKS

Grade	Cutoff
A A-	94% 90%
B+ B B-	87% 84% 80%
C+ C C-	77% 74% 70%
D+ D D-	67% 64% 60%
F	<60%

Course Outline

All instructions, assignments, readings, rubrics and essential information will be on the Canvas website at utexas.instructure.com. Check Canvas regularly. **Changes** to the schedule may be made at my discretion if circumstances require. I will announce any such changes in class and will also communicate them via a Canvas announcement. It is your responsibility to note these changes when announced, and I will do my best to ensure that you are notified of changes with as much advance notice as possible.

Week	Class Topic	Date Posted	Due dates
1.	Introduction		
2	Linear Algebra – review		
3	Essential Data science libraries-NumPy-Pandas		
4	Calculus and Optimization – review	PSET 1	
5	Probability – review		
6	Regression, Regularization, Dimension reduction	PSET 2	PSET 1
7	Finding Similar items		
8	Neural Networks		PSET 2
9	Probabilistic Learning		
10	Graphs and Graphical Models	PSET 3	
11	Paper discussions and presentation in class		
12	Late Midterm		PSET 3
13	Final project presentation		Final report due in class. Report is limited to 13-15 pages + 1-2 page of references.

Final report: 13-15 pages report + 1-2 page for references. Single line and 12 pt font. Follow a conference style paper.

[Cao] Data Science: A Comprehensive Overview, ACM Computing Surveys, Volume 50, Issue 3 Article No.: 43, pp 1–42.

[Leman Akoglu et al.] A Survey of Learning Causality with Data: Problems and Methods, ACM Computing Surveys, Volume 53, Issue 4, Article No.: 75, pp 1–37.

[Tang et al] Dimensionality Reduction Methods for Brain Imaging Data Analysis, ACM Computing Surveys, Volume 54, Issue 4, Article No.: 87, pp 1–36.

[Sun et al.] A Survey of Optimization Methods from a Machine Learning Perspective, https://arxiv.org/pdf/1906.06821.pdf

[Bennett and Parrado-Hernandez] The Interplay of Optimization and Machine Learning Research, Journal of Machine Learning Research, https://dl.acm.org/doi/pdf/10.5555/1248547.1248593

[Wu and Yao] Data Optimization in Deep Learning: A Survey

https://arxiv.org/abs/2310.16499

TEXAS
The University of Texas at Austin

[Gaurav Menghani], Efficient Deep Learning: A Survey on Making Deep Learning Models Smaller, Faster, and Better, ACM Computing Surveys, Volume 55, Issue 12, Article No.: 259, Pages 1 – 37, 2023, https://dl.acm.org/doi/10.1145/3578938

Course Policies and Disclosures

ACADEMIC INTEGRITY EXPECTATIONS

Students who violate University rules on academic dishonesty are subject to disciplinary penalties, including the possibility of failure in the course and/or dismissal from the University. Since such dishonesty harms the individual, all students, and the integrity of the University, policies on academic dishonesty will be strictly enforced. For further information, please visit the Student Conduct and Academic Integrity website at: http://deanofstudents.utexas.edu/conduct.

CONFIDENTIALITY OF CLASS RECORDINGS

Class recordings are reserved only for students in this class for educational purposes and are protected under FERPA. The recordings should not be shared outside the class in any form. Violation of this restriction by a student could lead to Student Misconduct proceedings.

GETTING HELP WITH TECHNOLOGY

Students needing help with technology in this course should contact the ITS Service Desk or [insert contact information for your local support unit(s) and for course materials, software, hardware, or other technology used in your course].

CONTENT WARNING

Our classroom provides an open space for the critical and civil exchange of ideas. Some readings and other content in this course will include topics that some students may find offensive and/or traumatizing. I'll aim to forewarn students about potentially disturbing content and I ask all students to help to create an atmosphere of mutual respect and sensitivity.

[Best practice discussions around content warnings also suggest including tags or other warnings on the Course Outline (above) next to the assigned material. Further discussion of content warning can be found at this page.]

BASIC NEEDS SECURITY

Any student who faces challenges securing their food or housing and believes this may affect their performance in the course is urged to contact the Dean of Students for support. UT maintains the UT Outpost, which is a free on-campus food pantry and career closet. Furthermore, if you are comfortable notifying me, please do so, as I may have additional resources I can share.

SHARING OF COURSE MATERIALS IS PROHIBITED

No materials used in this class, including, but not limited to, lecture hand-outs, videos, assessments (quizzes, exams, papers, projects, homework assignments), in-class materials, review sheets, and additional problem sets, may be shared online or with anyone outside of the class without explicit, written permission of the instructor. Unauthorized sharing of materials promotes cheating. The University is well aware of the sites used for sharing materials, and any materials found online that are associated with you, or any suspected unauthorized sharing of materials, will be reported to Student Conduct and Academic Integrity in the Office of the Dean of Students. These reports can result in sanctions, including failure of the course.

RELIGIOUS HOLY DAYS

By UT Austin policy, you must notify me of your pending absence as far in advance as possible of the date of observance of a religious holy day. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, you will be given an opportunity to complete the missed work within a reasonable time after the absence.

NAMES AND PRONOUNS

Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender identity & expression, and nationalities. Class rosters are provided to the instructor with the student's legal name, unless they have added a "chosen name" with the registrar's office, which you can do so here. I will gladly honor your request to address you by a name that is different from what appears on the official roster, and by the pronouns you use (she/he/they/ze, etc). Please advise me of any changes early in the semester so that I may make appropriate updates to my records. For instructions on how to add your pronouns to Canvas, visit this site. More resources available on the Gender and Sexuality Center's website, www.utgsc.org.

LAND ACKNOWLEDGMENT

I would like to acknowledge that we are meeting on the Indigenous lands of Turtle Island, the ancestral name for what now is called North America. Moreover, I would like to acknowledge the Alabama-Coushatta, Caddo, Carrizo/Comecrudo, Coahuiltecan, Comanche, Kickapoo, Lipan Apache, Tonkawa and Ysleta Del Sur Pueblo, and all the American Indian and Indigenous Peoples and communities who have been or have become a part of these lands and territories in Texas.

University Resources for Students

DISABILITY & ACCESS (D&A)

[This required syllabus content is repeated from above. It may be included in either place, or both.]
The university is committed to creating an accessible and inclusive learning environment consistent with university policy and federal and state law. Please let me know if you experience any barriers to learning so I can work with you to ensure you have equal opportunity to participate fully in this course. If you are a student with a disability, or think you may have a disability, and need accommodations please contact Disability & Access

(D&A). Please refer to the D&A website for more information: http://diversity.utexas.edu/disability/. If you are already registered with D&A, please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations and needs in this course.

COUNSELING AND MENTAL HEALTH CENTER (CMHC)

I urge students who are struggling for any reason and who believe that it might impact their performance in the course to reach out to me if they feel comfortable. This will allow me to provide any resources or accommodations that I can. If immediate mental health assistance is needed, call the Counseling and Mental Health Center (CMHC) at 512-471-3515 or you may also contact Bryce Moffett, LCSW (iSchool CARE counselor) at 512-232-2983. Outside CMHC business hours (8a.m.-5p.m., Monday-Friday), contact the CMHC 24/7 Crisis Line at 512-471-2255.

UNIVERSITY HEALTH SERVICES (UHS)

Your physical health and wellness are a priority. University Health Services is an on-campus high-quality medical facility providing care to all UT students. Services offered by UHS include general medicine, urgent care, a 24/7 nurse advice line, gynecology, sports medicine, physical therapy, lab and radiology services, COVID-19 testing and vaccinations and much more. For additional information, visit https://healthyhorns.utexas.edu or call 512-471-4955.

SANGER LEARNING CENTER

Did you know that more than one-third of UT undergraduate students use the Sanger Learning Center each year to improve their academic performance? All students are welcome to take advantage of Sanger Center's classes and workshops, private learning specialist appointments, peer academic coaching, and tutoring for more than 70 courses in 15 different subject areas. For more information, please visit https://ugs.utexas.edu/slc or call 512-471-3614 (JES A332)."

STUDENT EMERGENCY SERVICES (SES)

Student Emergency Services in the Office of the Dean of Students helps students and their families during difficult or emergency situations. Assistance includes outreach, advocacy, intervention, support, and referrals to relevant campus and community resources. If you need to be absent from class due to a family emergency, medical or mental health concern, or academic difficulty due to crisis or an emergency situation, you can work with Student Emergency Services. SES will document your situation and notify your professors. Additional information is available at https://deanofstudents.utexas.edu/emergency/ or by calling 512-471-5017.

Important Safety Information

If you have concerns about the safety or behavior of fellow students, TAs or professors, contact BCCAL (the Behavior Concerns and COVID-19 Advice Line) at https://safety.utexas.edu/behavior-concerns-advice-line or by calling 512-232-5050. Confidentiality will be maintained as much as possible, however the university may be required to release some information to appropriate parties.

CLASSROOM SAFETY AND COVID-19

- For any illness, students should stay home if they are sick or contagious, not only to stop the spread, but also to promote their personal wellness.
- The university will continue to provide rapid antigen self-test kits at distribution sites throughout campus. Students can receive up to four tests at a time.
- The university will provide symptomatic COVID-19 testing on campus for all students, faculty and staff.
- UHS maintains up-to-date resources on COVID, which can be found here:
    - COVID-19 Information and Resources
    - COVID-19 Exposure Action Chart

CARRYING OF HANDGUNS ON CAMPUS

Texas' Open Carry law expressly prohibits a licensed to carry (LTC) holder from carrying a handgun openly on the campus of an institution of higher education such as UT Austin. Students in this class should be aware of the following university policies:

- Students in this class who hold a license to carry are asked to review the university policy regarding campus carry.
- Individuals who hold a license to carry are eligible to carry a concealed handgun on campus, including in most outdoor areas, buildings and spaces that are accessible to the public, and in classrooms.
- It is the responsibility of concealed-carry license holders to carry their handguns on or about their person at all times while on campus. Open carry is NOT permitted, meaning that a license holder may not carry a partially or wholly visible handgun on campus premises or on any university driveway, street, sidewalk or walkway, parking lot, parking garage, or other parking area.
- Per my right, I prohibit carrying of handguns in my personal office. Note that this information will also be conveyed to all students verbally during the first week of class. This written notice is intended to reinforce the verbal notification, and is not a "legally effective" means of notification in its own right.

TITLE IX DISCLOSURE

[If this disclosure is included in the syllabus, the Title IX office has specificed the following wording.]
Beginning January 1, 2020, Texas Education Code, Section 51.252 (formerly known as Senate Bill 212) requires all employees of Texas universities, including faculty, to report any information to theTitle IX Office regarding sexual harassment, sexual assault, dating violence and stalking that is disclosed to them. Texas law requires that all employees who witness or receive any information of this type (including, but not limited to, writing assignments, class discussions, or one-on-one conversations, or third party reports) must be report it. Before talking with me, or with any faculty or staff member about a Title IX related incident, please remember that I will be required to report this information to the Title IX Coordinator. If you would like to speak with someone who can provide support or remedies without making an official report to the university, please email supportandresources@austin.utexas.edu. For more information about reporting options and resources, visit

http://www.titleix.utexas.edu/, contact the Title IX Office via email at titleix@austin.utexas.edu, or call 512-471-0419.

Although graduate teaching and research assistants are not subject to Texas Education Code, Section 51.252, they are still mandatory reporters under Federal Title IX laws and are required to report a wide range of behaviors we refer to as sexual misconduct, including the types of sexual misconduct covered under Texas Education Code, Section 51.252.The Title IX office has developed supportive ways to respond to a survivor and compiled campus resources to support all impacted by a Title IX incident.

CAMPUS SAFETY

The following are recommendations regarding emergency evacuation from the Office of Campus Safety and Security, 512-471-5767,

- Students should sign up for Campus Emergency Text Alerts at the page linked above.
- Occupants of buildings on The University of Texas at Austin campus must evacuate buildings when a fire alarm is activated. Alarm activation or announcement requires exiting and assembling outside.
- Familiarize yourself with all exit doors of each classroom and building you may occupy. Remember that the nearest exit door may not be the one you used when entering the building.
- Students requiring assistance in evacuation shall inform their instructor in writing during the first week of class.
- In the event of an evacuation, follow the instruction of faculty or class instructors. Do not re-enter a building unless given instructions by the following: Austin Fire Department, The University of Texas at Austin Police Department, or Fire Prevention Services office.
- For more information, please visit emergency preparedness.