

INFO5501: FUNDAMENTALS OF DATA ANALYTICS

Fall 2023

Instructor:	Junhua Ding, Ph.D.	Time:	Thursday: 5:30pm - 8:20pm
Email:	junhua.ding@unt.edu	Place:	NTDP K150
Office:	Discovery Park E292G	Phone:	(940)565-2186

Teaching Assistants

1. Huyen T.N. Nguyen, huyennguyen5@my.unt.edu
2. Lavanya Pobbathi, lavanyapobbathi@my.unt.edu
3. Sai Kolapudi, saikolapudi@my.unt.edu

Course Pages

- Course management: Canvas: <https://unt.instructure.com/>

Office Hours

- Thursday: 12:30pm - 4:30pm, or by appointment.

Textbooks

- No required textbook. The online materials on *Fundamentals of Data Science*, by Stephen Elston, at [github.com](https://github.com/StephenElston/CSCI-E-83) are useful for understanding the content in this course. The materials can be found at: <https://github.com/StephenElston/CSCI-E-83>

Prerequisite

1. General programming experience is assumed. The Python programming language will be used for this class and reviewed as appropriate.
2. Basic linear algebra and probability knowledge is needed.

Software

The following software and computing environments might be used in this course.

- Google Colab is needed. You may use Python Jupyter Notebook, JupyterLab, Python Frameworks such as NumPy, SciPy, Pandas in Google Colab.

Objectives

This course teaches mathematical and statistical fundamentals for data science and focuses on the modern computational statistical modeling methods. The course covers basic computational statistical inference, the properties and behavior of linear models, conditional probability, Bayesian models, inference by resampling and time series methods. These skills and knowledge enable students to understand data science literature and principles behind data science tools and algorithms, and provide students a solid foundation for data analytics, data science modeling, and machine learning. The objectives of this course:

1. Be able to explore big data sets to gain valuable insights.
2. Be able to use computational statistical methods to build data science solutions.
3. Be able to explain neural networks (deep and otherwise) compare to other machine learning models.

4. Be able to use linear models and resampling methods in data science practice.
5. Be able to conduct inference with Bayesian models, and explain the relationship with frequentist models.
6. Be able to identify and model data with time dependency and build forecasts.

Topics

The following is the tentative schedule with the covered topics. Actual schedule may be adjusted according to progress:

1. Introduction to Data Science
2. Conditional Probability
3. Sampling and Simulation
4. Parameter Estimation and Likelihood
5. Confidence Intervals and Bootstrap Resampling
6. Bayes Models
7. Linear Models
8. Models in High Dimensions
9. Generalizing the Linear Model
10. Forecasting and Time Series Models
11. Markov Chain Monte Carlo
12. Least Squares Optimization

Grading Policy

- Students are required to attend the class on time, complete all assignments, quizzes, exams, and readings on time. Grading will be based on assignments to be assigned as the course proceeds.
Grades will be computed as: 1. Assignments: 20%, 2. 2 quizzes: 20%, 3. 2 Exams: 60%.
- **Grading Scale:** A: 90-100; B: 80-89; C: 70-79; D: 60-69; F: 59 or below.

All assignments are post in Canvas, and students are required to upload their work into Canvas. The final grade is calculated based on grade points of assignments, quizzes, and exams. All quizzes and exams are close-book, close-note exams to be given in the classroom during class meeting time.

Incompletes

A grade of incomplete (I) will be given only for a justifiable reason (such as a serious illness or military service) and only if you are passing the course. It is your responsibility to contact the instructor to request an incomplete and discuss requirements for completing the course. If you do not remove the incomplete within the timeframe agreed upon with the instructor or within one calendar year, you will receive a grade of an F. Please refer to <http://essc.unt.edu/registrar/academic-record-incomplete.html> for more information.

Withdrawal

A grade of withdraw (W) or withdraw-failing (WF) will be given depending on your participation and grades to date. If you simply disappear and do not file a formal UNT withdrawal form, you may receive a grade of an F.

Teaching Philosophy

This course will be taught in traditional lectures although some of the lectures are delivered via video conference system Zoom. Attending the class on time, and on-class discussions are strongly encouraged. Although software tools such as Python Notebook, and TensorFlow will be needed to complete the assignments, few lectures will be given on how to use the tools. The course employs a combination of theory and hands-on experience using Python programming tools.

This course employs lecture capture technology to record class sessions if the lectures are delivered via a video conference system. 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.

Technical Assistance

UIT Help Desk: <http://www.unt.edu/helpdesk/index.htm>. The University of North Texas provides student technical support in the use of Zoom and Canvas and supported resources. The student help desk may be reached at:

- Email: helpdesk@unt.edu
- Phone: 940.565-2324
- In Person: Sage Hall, Room 130

Hours

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

Canvas technical requirements: <https://clear.unt.edu/supported-technologies/canvas/requirements>

Minimal Technical Skills and Resources Needed

Each student should be able to access computers that have environment for running and developing Python programs with databases Microsoft Excel and mySQL. Google Colab (<https://colab.research.google.com/>) is a recommended environment for developing course projects and running sample codes from the textbook.

Communication

Students can email their questions to the instructor and the teacher assistant (TA). They are also encouraged to talk to the instructor and TA during the office hours. Emails are normally respond within 24 hours, and all assignments, quizzes, projects, and papers should be graded within 10 days after the submission deadline.

Class Policy

- **Diversity and Inclusion Statement:** I desire for students from all diverse backgrounds and perspectives be encouraged by this course, that learning needs are addressed, and that the diversity you carry into class is seen as a resource, strength and benefit. We should all be respectful of diversity: gender, sexuality, disability, age, socioeconomic status, ethnicity, race, culture, and more. All of your voices should be heard. Suggestions are welcomed and respected – please let me know ways to improve the course.
- **Prohibition of Discrimination, Harassment, and Retaliation:** 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. According UNT Policy 16.004, UNT prohibits discrimination and harassment because of race, color, national origin, religion, sex, sexual orientation, gender identity, gender expression, age, disability, genetic information, veteran status, or any other characteristic protected under applicable federal or state law in its application and admission processes; educational programs and activities; employment policies, procedures, and processes; and university facilities. The University takes active measures to prevent such conduct and investigates and takes remedial action when appropriate.

- **Attendance Policy:** You are expected to attend class via Zoom. You are responsible for announcements and assignments given in class. If you miss a class, it is up to you to obtain notes and any other information that was provided in the class. Those who do not attend class or review the recorded lectures in a timely manner can count on doing poorly in this course.
- **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.

Students caught cheating or plagiarizing will receive a “0” for that particular assignment or exam (or specify alternative sanction, such as course failure). Additionally, the incident will be reported to the Office of Student Rights and Responsibilities, which may impose for further penalty. According to the UNT catalog, the term “cheating” includes, but is not limited to: (a). use of any unauthorized assistance in taking quizzes, tests, or examinations; (b). dependence upon the aid of sources beyond those authorized by the instructor in writing papers, preparing reports, solving problems, or carrying out other assignments; (c). the acquisition, without permission, of tests or other academic material belonging to a faculty or staff member of the university; (d). dual submission of a paper or project, or resubmission of a paper or project to a different class without express permission from the instructor(s); or (e). any other act designed to give a student an unfair advantage. The term “plagiarism” includes, but is not limited to: (a). the knowing or negligent use by paraphrase or direct quotation of the published or unpublished work of another person without full and clear acknowledgment; and (b). the knowing or negligent unacknowledged use of materials prepared by another person or agency engaged in the selling of term papers or other academic materials.

The Department of Information Science (DIS) Academic Integrity and Misconduct Guidelines on the website at <https://informationscience.unt.edu/departamental-guidelines> have the following on page 6, which are also applied to this course:

“DIS instructors will impose the following academic penalties for violations of the DIS Academic Integrity and Misconduct Policy:

First Offense: 0 for the assignment ; if a Practicum site, site denied. Second Offense: F (fail) for the course, entered in the gradebook; Practicum application deferred for one semester.

Per UNT Policy 06.003, the course instructor retains the right to determine specific sanctions for their course and to set additional policies and procedures that do not conflict with DIS or UNT policies.”

- **ADA Policy:** The University of North Texas makes reasonable academic accommodation for students with disabilities. Students seeking reasonable accommodation must first register with the Office of Disability Accommodation (ODA) to verify their eligibility. If a disability is verified, the ODA will provide you with a reasonable accommodation letter to be delivered to faculty to begin a private discussion regarding your specific needs in a course. You may request reasonable accommodations at any time, however, ODA notices of reasonable 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 reasonable accommodation for every semester and must meet with each faculty member prior to implementation in each class. Students are strongly encouraged to deliver letters of reasonable accommodation during faculty office hours or by appointment. Faculty members have the authority to ask students to discuss such letters during their designated office hours to protect the privacy of the student. For additional information see the Office of Disability Accommodation website at <http://www.unt.edu/oda>. You may also contact them by phone at 940.565.4323.
- **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.

- **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 Center for Student Rights and Responsibilities 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. The Code of Student Conduct can be found at deanofstudents.unt.edu/conduct.
- **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: eagleconnect.unt.edu/
- **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 [insert administration dates] 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 at <http://spot.unt.edu/> or email: spot@unt.edu.
- **Sexual Assault Prevention:** UNT is committed to providing a safe learning environment free of all forms of sexual misconduct, including sexual harassment sexual assault, domestic violence, dating violence, and stalking. Federal laws (Title IX and the Violence Against Women Act) and UNT policies prohibit discrimination on the basis of sex, and therefore prohibit sexual misconduct. If you or someone you know is experiencing sexual harassment, relationship violence, stalking, and/or sexual assault, there are campus resources available to provide support and assistance. UNT's Survivor Advocates can assist a student who has been impacted by violence by filing protective orders, completing crime victim's compensation applications, contacting professors for absences related to an assault, working with housing to facilitate a room change where appropriate, and connecting students to other resources available both on and off campus. The Survivor Advocates can be reached at SurvivorAdvocate@unt.edu or by calling the Dean of Students Office at 940-565- 2648. Additionally, alleged sexual misconduct can be non-confidentially reported to the Title IX Coordinator at oeo@unt.edu or at (940) 565 2759.