CSCE 5300.004 – Introduction to Big Data and Data Science

Instructor: Pavlo Tymoshchuk
Office: E245Q
Office hours: 4:00 PM - 5:00 PM, TuTh;

Semester: Spring 2024
Time: 11:30 AM - 2:20 PM, Sa
Place: NTDP E265
Pavlo.Tymoshchuk@unt.edu
https://unt.zoom.us/j/7873105

Teaching Assistants
Katlaganti, Sravani, sravanikatlaganti@my.unt.edu
Office Hours: Monday, 4:00 PM - 6:00 PM
Zoom: https://unt.zoom.us/j/83657130402
Yerramreddy, Hemanth, sriramrelangi@my.unt.edu
Office Hours: We: 4:00 PM - 5:30 PM, Sa: 3:00 PM - 4:30 PM
Zoom: https://zoom.us/j/7256112597?pwd=bjhJblZTU0JCY1FXemVTZ3Y1UzBJUT09

Course Description
Introduction to big data and data science including an overview of the field, technical challenges, big data processing systems, structured and unstructured data processing, data analytics and learning, data visualization, statistical computational approaches, empirical methods in computer science, practical applications, privacy and ethics. Emphasis on big data and its effect on other topics within data science, its technical characteristics, and state-of-the-art big data analytics architectures and tools.

Learning Outcomes
By the end of the course you will
– Be familiar with technical challenges of the field
– Be familiar with computational approaches used to process big data and in data science
– Be familiar with practical applications of big data and data science methods
– Be able to process structured and unstructured data
– Be able to use empirical methods in computer science
– Have a fundamental understanding of data analytics and learning methods
– Be able to visualize data
– Be familiar with privacy and ethics of the field

Course Requirements
Attendance: attendance is strongly recommended and encouraged. Lectures and classes participation will contain vital information needed to do well on exams.

Textbooks

**Schedule**

<table>
<thead>
<tr>
<th>Week</th>
<th>Topic</th>
<th>References</th>
<th>Assignment Outcomes</th>
</tr>
</thead>
</table>

Week 8, March 9 - midterm exam to be held in the classroom in closed-book mode (pen and A4 papers/laptop only, no Internet, no mobile phones, no talking with neighbors)
No midterm exam rescheduling/transfer, no additional midterm exam. The date and time of the midterm exam can be only changed by Registrar or CSE department due to scheduling conflicts with other courses

Week 9
Spring break (no classes)


Week 17, May 4 – final exam (according to https://registrar.unt.edu/exams/final-exam-schedule/fall.html)
The date and time of the final exam can be only changed by UNT Registrar and CSE department due to scheduling conflicts with other courses
Final exam to be held in the classroom in closed-book mode (pen and A4 papers/laptop only, no Internet, no mobile phones, no talking with neighbors)

**Evaluation**
Exams I and II, 15% and 20% correspondingly
Assignments, 35%
Final Project, 15%
10% penalty from maximal grade per 1 day delay (0<delay<=24h) for late submission and 0 grade for submission after 4th day apply
Activity (Active Learning/Hands-On Learning/Quizzes/Answers in the classroom), 15%. No activity rescheduling/transfer
Attendance, 5%
A: 90-100%; B: 80-89%; C: 70-79%; D: 60-69%; F: <60%
Remarks:
- assignments and project are returned for correction if their plagiarism level exceeds 30% threshold;
- students are welcome to contact instructor/TAs if they feel that assignments/active learning tasks are too complex/easy for them to make corresponding adjustment.
Prerequisites: Probability theory, Calculus, Data structures, familiarity with Mallab/Python/R/C/C++/Java/MySQL. Note that Python will mainly be used in the programming assignments of this course.

Course Policies and Procedures
1. Emails will be answered as promptly as possible. Emails outside normal working hours (8 am to 5 pm) will be answered on the next working day.
2. Work handed in for grade (homework, project report, etc.) MUST BE YOUR OWN effort only. Students are NOT allowed to use online solutions from previous course offerings, websites, etc. This will be checked on plagiarism. The students should adhere to the UNT policies and procedures on the Code of Academic Integrity. 06.003 Student Academic Integrity https://policy.unt.edu/policy/06-003. Plagiarism WILL result in a score of 0 for the assessment in which it occurs.
3. Any appeals on grading must be made online in Canvas (not by email) no later than two days after the grading is complete. This applies to quizzes, assignments, project, midterm exam and final exam. Appeals should be justified, specific and contain a valid reason. Appeals such, e.g., as “I had poor Internet connection before deadline”, “My grade is close to a border with A/B/C grade”, “I am on probation”, “I believe I deserve more for this assessment”, “My overall grade has reduced as a result of this assessment”, “I was absent in the class but I do not have written confirmation of solid reason of my absence” or appeal made not in online in Canvas or appeal made later than two days after the grading is complete are NOT valid and such appeals are not processed. Grading is done carefully, and, in most cases, appeals do not result in a change of mark.
4. 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 classrooms, labs, discussion groups, field trips, etc. The Code of Student Conduct can be found at https://policy.unt.edu/policy/07-012
5. Usage of cell phones, earphones, and other electronic devices, or recording of lectures is prohibited. Usage of laptops and tablets is permitted for class purposes, only after obtaining permission from the instructor. Usage of classrooms computers, if any, are not allowed, while the class is in session. Any student who uses an unauthorized device will lose 1 point (out of 100) and may be asked to leave the classroom.

Disabilities Accommodation:
The University of North Texas complies with Section 504 of the 1973 Rehabilitation Act and with the Americans with Disabilities Act of 1990. The University of North Texas provides academic adjustments and auxiliary aids to individuals with disabilities, as defined under the law. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you believe you have a disability requiring
accommodation, please contact the instructor and/or the Office of Disability Accommodation at 940-565-4323 during the first week of class.

Course Policies:
● ABSOLUTELY NO LATE work will be graded unless specific arrangements are made with the instructor in advance. The University is committed to providing a reliable online course system to all users. However, 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 and TA and contact the UNT Student Help Desk: helpdesk@unt.edu or 940.565.2324 and obtain a ticket number. The instructor, TA, and the UNT Student Help Desk will work with the student to resolve any issues at the earliest possible time.
● All assignments will be turned in by the designated date due. Assignments may be submitted on Canvas in the appropriate drop box unless otherwise indicated.
● ALL requests for extensions on assignments must be made prior to the due date, in person, and must be for a valid “emergency” reason. In extreme circumstances, contact after the due date may be accepted if there is a COMPELLING reason.
● Attendance and participation is required, it is a part of your grade, and will be monitored in order to ensure that all groups operate at peak efficiency. You are responsible for all discussion, lecture and other information disseminated during the lecture period, regardless of whether you attend or not. You are also responsible for all team assignments made by your team lead and deliverable leads regardless of your attendance. You must provide documentation for excused absences for emergencies etc.
● Lectures and assignments are included in Canvas. However, you should regularly check the class website, as well as take note of in-class announcements for changes in the schedule or assignments.
● You should plan to spend, on average, about 10-15 hours per week outside of the normal class meetings working on the various aspects of your project. As deadlines draw near, the time commitment may increase.

Professional Conduct and Ethical Responsibility:
You will be representing yourself, your team, the CSE Department and UNT as you conduct the activities needed to deliver your capstone project. You are expected to conduct yourself professionally during team, class and sponsor interactions both verbally and in writing. This responsibility carries through in the project as well. You are learning new skills and applying what you have learned in other courses. The ACM Code of Ethics, available at: https://ethics.acm.org/, reflects the expectations of your conduct in this course. You are also expected to abide by the UNT Code of Conduct.

Collaboration and Cheating:
Collaboration among students in class is most certainly encouraged, as it is believed that it provides a better learning environment and is required for team assignments. All resources used should be clearly cited in written work of any kind, both individual and team. For further details and clarifications regarding collaboration and cheating, view the university Student Rights and Responsibilities web page.

Student Perceptions of Teaching (SPOT)
The Student Perceptions of Teaching (SPOT) is a requirement for all organized classes at UNT. This short survey will be made available to you at the end of the semester, providing you a chance to comment on how this class is taught. I am very interested in the feedback I get from students, as I work to continuously improve student learning quality and my teaching and satisfy student needs. I consider the SPOT to be an important part of your participation in this class since our goals in the scope of this class are your good knowledge and grades. SPOT will be available in each time of the term especially at its start.

Course Technology & Skills
Minimum Technology Requirements
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. There a Student Help Desk is at UNT 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)
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).

Rules of Engagement
Rules of engagement refer to the way students are expected to interact with each other and with their instructors. Here are some general guidelines:
- While the freedom to express yourself is a fundamental human right, any communication that utilizes cruel and derogatory language based on 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 will not be tolerated.
- Treat your instructor and classmates with respect in any communication online or face-to-face, even when their opinion differs from your own.
- Ask for and use the correct name and pronouns for your instructor and classmates.
- Speak from personal experiences. Use “I” statements to share thoughts and feelings. Try not to speak on behalf of groups or other individual’s experiences.
- Use your critical thinking skills to challenge other people’s ideas, instead of attacking individuals.
- Avoid using all caps while communicating digitally. This may be interpreted as “YELLING!”
- Be cautious when using humor and do not use sarcasm in emails or discussion posts as sense can be difficult to interpret digitally.
- Avoid using “text-talk” unless explicitly permitted by your instructor.
- Proofread and fact-check your sources.
- Keep in mind that online posts can be permanent, so think first before you type.
See Engagement Guidelines (https://clear.unt.edu/online-communication-tips) for more information.
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 with concerns about COVID-19. Please contact instructor 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 instructor prior to being absent so instructor can decide 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), 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 has a higher priority.

Class Materials for Remote Instruction
Remote instruction may be necessary if community health conditions change, or you need to self-isolate or quarantine due to COVID-19. Students will need access to a webcam and microphone to participate in fully remote portions of the class. Information on how to be successful in a remote learning environment can be found at https://online.unt.edu/learn.

Face Covering
Consistent with CDC Guidelines and state policy, individuals are not required to wear face coverings on UNT’s campuses. While non-vaccinated individuals are encouraged to wear a face covering, all students and faculty are welcome to wear one in class or on campus if they desire.

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. We emphasize that individual work such as homework assignments, quizzes and exams must be done on your own and that cheating will result in automatic failure of the course. In addition, the incident will be reported to the CSE department and Academic Integrity department, in accordance with the UNT policy. Do not discuss solutions or share copies of individual work. Unacceptable collaboration will be considered a violation of a Code of Student Conduct and will result in a failing grade for the course. Please make sure if you obtain ideas from others, you refer your source. If questions arise during working on a problem, please feel free to contact the instructor for a clarification.

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/).

Prohibition of Discrimination, Harassment, and Retaliation (Policy 16.004)
The University of North Texas (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.

**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: [https://my.unt.edu](https://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 [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 ([http://spot.unt.edu/](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 based on 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.

Important Notice for F-1 Students taking Distance Education Courses

Federal Regulation
To read detailed Immigration and Customs Enforcement regulations for F-1 students taking online courses, please go to the Electronic Code of Federal Regulations website (http://www.ecfr.gov/). The specific portion concerning distance education courses is located at Title 8 CFR 214.2 Paragraph (f)(6)(i)(G).
The paragraph reads:
(G) For F-1 students enrolled in classes for credit or classroom hours, no more than the equivalent of one class or three credits per session, term, semester, trimester, or quarter may be counted toward the full course of study requirement if the class is taken on-line or through distance education and does not require the student's physical attendance for classes, examination or other purposes integral to completion of the class. An on-line or distance education course is a course that is offered principally using television, audio, or computer transmission including open broadcast, closed circuit, cable, microwave, or satellite, audio conferencing, or computer conferencing. If the F-1 student’s course of study is in a language study program, no on-line or distance education classes may be considered to count toward a student's full course of study requirement.

University of North Texas Compliance
To comply with immigration regulations, an F-1 visa holder within the United States may need to engage in an on-campus experiential component for this course. This component (which must be approved in advance by the instructor) can include activities such as taking an on-campus exam, participating in an on-campus lecture or lab activity, or other on-campus experience integral to the completion of this course. If such an on-campus activity is required, it is the student’s responsibility to do the following:
(1) Submit a written request to the instructor for an on-campus experiential component within one week of the start of the course.
(2) Ensure that the activity on campus takes place and the instructor documents it in writing with a notice sent to the International Student and Scholar Services Office. ISSS has a form available that you may use for this purpose.
Because the decision may have serious immigration consequences, if an F-1 student is unsure about his/her need to participate in an on-campus experiential component for this course, he/she should contact the UNT International Student and Scholar Services Office (telephone 940-565-2195 or email internationaladvising@unt.edu) to get clarification before the one-week deadline.

Student Verification
UNT takes measures to protect the integrity of educational credentials awarded to students enrolled in distance education courses by verifying student identity, protecting student privacy, and notifying students of any special meeting times/locations or additional charges associated with student identity verification in distance education courses.
See UNT Policy 07-002 Student Identity Verification, Privacy, and Notification and Distance Education Courses (https://policy.unt.edu/policy/07-002).

Use of Student Work
A student owns the copyright for all work (e.g., software, photographs, reports, presentations, and email postings) he/she creates within a class and the University is not entitled to use any student work without the student’s permission unless all the following criteria are met:
● The work is used only once.
● The work is not used in its entirety.
Use of the work does not affect any potential profits from the work.
The student is not identified.
The work is identified as student work.

If the use of the work does not meet all the above criteria, then the University office or department using the work must obtain the student’s written permission.
Download the UNT System Permission, Waiver and Release Form

Transmission and Recording of Student Images in Electronically Delivered Courses
1. No permission is needed from a student for his or her image or voice to be transmitted live via videoconference or streaming media, but all students should be informed when courses are to be conducted using either method of delivery.
2. In the event an instructor records student presentations, he/she must obtain permission from the student using a signed release to use the recording for future classes in accordance with the Use of Student-Created Work guidelines above.
3. Instructors who video-record their class lectures with the intention of re-using some or all of recordings for future class offerings must notify students on the course syllabus if students' images may appear on video. Instructors are also advised to provide accommodation for students who do not wish to appear in class recordings.

Example: 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.

No notification is needed if only audio and slide capture is used or if the video only records the instructor's image. However, the instructor is encouraged to let students know the recordings will be available to them for study purposes.

Class Recordings & Student Likenesses
Synchronous (live) sessions in this course can 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.

Academic Support & Student Services
Student Support Services
Mental Health
UNT provides mental health resources to students to help ensure there are numerous outlets to turn to that wholeheartedly care for and are there for students in need, regardless of the nature of an issue or its severity. Listed below are several resources on campus that can support your academic success and mental well-being:

- Student Health and Wellness Center (https://studentaffairs.unt.edu/student-health-and-wellness-center)
- Counseling and Testing Services (https://studentaffairs.unt.edu/counseling-and-testing-services)
- UNT Care Team (https://studentaffairs.unt.edu/care)
- Individual Counseling (https://studentaffairs.unt.edu/counseling-and-testing-services/services/individual-counseling)
**Chosen Names**

A chosen name is a name that a person goes by that may or may not match their legal name. If you have a chosen name that is different from your legal name and would like that to be used in class, please let the instructor know. Below is a list of resources for updating your chosen name at UNT.

- UNT Records
- UNT ID Card
- UNT Email Address
- Legal Name

*UNT euIDs cannot be changed at this time. The collaborating offices are working on a process to make this option accessible to UNT community members.*

**Pronouns**

Pronouns (she/her, they/them, he/him, etc.) are a public way for people to address you, much like your name, and can be shared with a name when making an introduction, both virtually and in-person. Just as we ask and don’t assume someone’s name, we should also ask and not assume someone’s pronouns. You can add your pronouns to your Canvas account so that they follow your name when posting to discussion boards, submitting assignments, etc.

Below is a list of additional resources regarding pronouns and their usage:

- What are pronouns and why are they important?
- How do I use pronouns?
- How do I share my pronouns?
- How do I ask for another person’s pronouns?
- How do I correct myself or others when the wrong pronoun is used?

**Additional 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)
- Pride Alliance (https://edo.unt.edu/pridealliance)
- UNT Food Pantry (https://deanofstudents.unt.edu/resources/food-pantry)

**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/)