LING 5405: Python for Text

Fall 2020

Monday/Wednesday, 4:30-5:50pm, remote meetings via Zoom (link on Canvas)

Professor: Dr. Alexis Palmer, she/her, DP B201A, Alexis.Palmer@unt.edu, 940-369-8079 (voicemail)
TA/Grader: Mayank Pinna, PinnaPinna@my.unt.edu

Office hours: Mondays 6-7pm, Tuesdays 3-6pm, please make an appointment using calendly – email me if you need an appointment at a different day and/or time

================================ LOGISTICS ================================================

Communication expectations:
1. Canvas – Canvas is the official site for course announcements from me to you. Please check it daily.
2. Email – If you have questions or concerns, email is the best way to reach me. Feel free to email me at any time (alexis.palmer@unt.edu) – I try to respond to emails within 24 hours, except on weekends.
3. Feedback – Feedback on assignments will be given through Canvas or through comments on Colab. I will do my best to give feedback on all assignments within one week of the date they were submitted. This timeframe may change as the circumstances of the semester change.
4. Grading – Grades will be posted on Canvas. I will do my best to grade all assignments within one week of the date they were submitted. This timeframe may change as the circumstances of the semester change.

Course structure:
This course will be meeting remotely, with the possibility of doing everything asynchronously. Live sessions will be held via Zoom and recorded. I strongly prefer that you attend the live sessions (at the scheduled class meeting times) if at all possible, but I do understand that schedules are especially unpredictable right now, and some students may have trouble attending class meetings.

Synchronous meetings will take place at 4:30 pm each Monday and Wednesday and will end no later than 5:50pm (usually earlier, I expect). On Mondays we will go over previous assignments and then have (optional) lab time for questions and discussions. On Wednesdays, we will begin with student presentations and then have (optional) lab time for questions and discussions. It’s all a bit experimental – this schedule may adjust as we go through the semester.

Class Materials for Remote Instruction
Students will need access to a computer or other device with a webcam and microphone, as well as reliable internet access, to participate in the 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. Here you can also find information about computer lab access. More information on technical support is available later in the syllabus.

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 contact the UNT Student Help Desk: helpdesk@unt.edu or 940.565.2324 and obtain a ticket number. The instructor and the UNT Student Help Desk will work with the student to resolve any issues at the earliest possible time.
Course Description and learning objectives:
Many disciplines increasingly deal with large amounts of textual data. The Python programming language is an easily-accessible tool which provides a foundation for learning to deal with textual data faster, more efficiently, and more systematically. With some basic knowledge of Python, scholars can easily access existing toolkits for text processing, data management and manipulation, and statistical analysis. This course is a friendly way to learn beginning to intermediate Python programming, starting from an assumption of zero programming knowledge or experience. Emphasis is placed on developing practical skills in analyzing text. Each student will execute a final project performing text analysis on data from their field of study.

Prerequisites: None.

Our course goals include the following:
- Understand Python’s data types and data structures
- Learn to write Python scripts for data processing and manipulation
- Learn to use existing Python modules and toolkits for text-based research (e.g. modules for working with Twitter data; the Natural Language Toolkit; etc.)
- Learn to use Python tools for interacting with digitally-available text corpora
- Become familiar with methods for representing text computationally
- Apply the learned methods for text analysis in a particular field

Course readings: All readings will be made available through Canvas.

Course credit requirements and grading
1. 25% - Weekly labs (14 labs, lowest 2 scores will be dropped)
2. 40% - Homework assignments (5 regular, 2 project-related)
3. 25% - Final project and write-up
4. 10% - Presentations

Discussion boards
Please use these to ask (and answer) questions, explore topic ideas, share new toolkits, etc.

Attendance & Participation
Attendance is STRONGLY PREFERRED but not strictly required. The course is designed to allow for asynchronous participation. The Zoom meetings will be full of useful and helpful information, though, so you’d be missing out! (They will also be recorded, but I hope you can join in person.)

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) 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.

What to expect:
The course is grouped into six focus areas: News as text, Literature as text, Web and social media as text, Representing text, Text classification, and Texts in many languages. Each week we will address a new topic, using the indicated types of data as our focus.

On Monday each week you will receive a new lab/tutorial, which will be a combination of instructional material, examples, and links to readings and/or videos. In addition, the lab will contain questions to be answered and problems to be solved. Your answers and solutions must be submitted no later than midnight on the following Sunday. At the start of Monday’s class, we will go over solutions for the labs just submitted – this means NO LATE LABS will be accepted.

LING 5405 students will give presentations on Wednesdays on topics related to the current week’s themes. Each presentation should last about 10 minutes, and each student will do from 1-3 presentations.

Detailed schedule can be found on Canvas. Usually:
- The lab and readings for each week will be released on Monday and will be due by midnight Sunday
- Monday’s class is for reviewing assignments, followed by lab time for questions and discussion
- Wednesday’s class is for student presentations, followed by lab time for questions and discussion
- Homework assignments are also due on Sundays – these will be assigned 2 weeks before they are due

Final project and write-up:
Working in groups of 2-3, students will complete a final programming and/or analysis project on a topic of interest, focused on textual data. I will help you to select projects that can be done in the short time frame. In addition to the implementation, each group will give a presentation and submit a written technical report.

Late assignments:
Labs may not be submitted late, but the lowest 2 grades for each will be dropped, giving you the flexibility to miss a couple over the course of the semester without penalty.

Homework assignments may not be submitted late, unless you have extenuating circumstances. In that case, you *must* contact me as soon as you know there will be a problem. For project-related work, I will consider extensions by (email) request. The sooner you submit your work, though, the sooner you’ll get feedback.

================================ ADDITIONAL INFORMATION ==================================

Linguistics Colloquia and COI Discovery Series Talks: The Department of Linguistics and the College of Information each offer a series of talks from experts in the field. Please plan to attend these if at all possible – I’ll be sure to announce the talks, and they will also appear on the course schedule.

Class behavior (University policy): 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 www.unt.edu/csrr

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.

Rules of Engagement
This is a discussion-based class, and our discussions may sometimes take us into areas where people disagree. Rules of engagement are some guidelines for how we are expected to interact with each other. Here are some general guidelines:

• While the freedom to express yourself is a fundamental human right, any communication that utilizes cruel and derogatory language on the basis 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 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 or sarcasm in emails or discussion posts as tone 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 these Engagement Guidelines (https://clear.unt.edu/online-communication-tips) for more information.

Academic Integrity: Please make sure that all the work you submit in this course is your own!!!
It is your responsibility to know the definitions and consequences of plagiarism, cheating etc. as outlined in the Code of Student Conduct and Discipline, available in the Undergraduate Catalog and online at http://catalog.unt.edu/content.php?catoid=5&navoid=244#Student_Standards_of_Academic_Integrity

Disability Accommodation Statement: “The University of North Texas 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 you with an accommodation letter to be delivered to faculty to begin a private discussion regarding your specific needs in a course. You 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 Office of Disability Accommodation website at http://www.unt.edu/oda. You may also contact them by phone at 940.565.4323.”
Students with any special needs or disabilities (including learning disabilities) should inform me of their needs as soon as possible so that proper accommodations can be made.

**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.

**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)
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)

**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.