INFO 5717

Networked Data Modeling and Processing

Section: 001, 005

SYLLABUS

Fall 2019

Table of Contents

COURSE INFORMATION
- INFO 5717, Sections 001 and 005, 3 Credit Hours
- Title: Networked Data Modeling and Processing
- **Online Meeting Dates**: See Table 2
- Online Meeting Time: 7:00 pm – 9:00 pm.

Instructor Contact Information

- **Dr. Jiangping Chen**, Professor. Department of Information Science, College of Information, University of North Texas
- Office: DP IS Department Chair Office E290BA
- Phone: (940) 369-5825
- Email address: jiangping.chen@unt.edu
- Office hours: Thursday 11:00am to 1:00pm or By appointment -- arrange via course messages

Teaching Assistant and Tutor

- **Haihua Chen,** Information Science Ph.D. student. Department of Information Science, College of Information, University of North Texas.
- Office: DP E292J
- Phone: (940) 220-0057
- Email address: haihua.chen@unt.edu
- Office hours: Wednesday 11:00am – 2:00pm.

Communicating with Your Instructor
This course will have a website in UNT Canvas (https://unt.instructure.com/login/canvas) for online discussion, assignment submissions, and sharing of reading materials. Students are welcome to make an appointment with the instructor and/or the teaching assistant (TA) to discuss course-related questions (online). If you need to schedule an individual online meeting with the instructor or the TA, please send her/him an email via the course website in Canvas Course Messages to make an appointment.

Course Pre-requisites, Co-requisites, and/or Other Restrictions

- Pre-requisite: INFO 5707: Data Modeling for Information Professionals, or by the consent of the instructor

Course Format

INFO 5717, Sections 001 and 005 hold online lectures by the instructor. The course uses Canvas, UNT's new learning management system. ALL course materials will be available at the course site on Canvas that is accessible to all students. And students will submit all assignments through the tools available on Canvas.

Course Description

Designed to meet the needs of data modeling, analysis, presentation, and access on the Internet and other networked environments. Focuses on issues relating to design and implementation of database-driven web systems. Students develop a thorough theoretical understanding of such systems and related issues, and obtain hands-on experience with data collecting, modeling, integration, and retrieval through working on a semester-long team project.

Course Goals, Learning Objectives

- Master basic concepts and components of a computer program
- Write Python programs to process and visualize data and data sets.
- Demonstrate the ability of crawling data from different resources on the web by using Python Program.
- Clean and preprocess raw text data using basic natural language processing techniques
- Design and implement applications that process, manage, and analyze text data in databases.
- Document and report on information processing and applications.

Materials

Textbook information (required):
1. Downey, Allen B. (2016). Think Python: How to Think Like a Computer Scientist, 2nd Edition. O'Reilly, ISBN-13: 978-1-491-93936-9.

2. Ryan Mitchell. (2018). Web Scraping with Python: Collecting More Data from the Modern Web, 2nd Edition. O'Reilly, ISBN-13: 978-1491985571.

3. Igor Milovanovic, Dimitry Foures, Giuseppe Vettigli. (2015). Python Data Visualization Cookbook, 2nd Edition.

Supplementary materials and/or readings (recommended):
4. Python Documentation: https://www.python.org/doc/.

Teaching Philosophy

The instructor will take a problem-solving approach and work together with student to understand how computer programs work. We will learn how to solve practical data processing, data management, and data visualization problems. She will monitor the progress of students and is open to suggestions from students. Students are expected to study 12-15 hours per week, and to submit their assignments on time to achieve satisfactory class performance. Interaction between the student and the instructor/TA is guaranteed and strongly encouraged. Students who don't have knowledge and experience in database are expected to spend extra hours on this course.

TECHNICAL REQUIREMENTS/ASSISTANCE

UIT Help Desk: http://www.unt.edu/helpdesk/index.htm

The University of North Texas provides student technical support in the use of 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 330

Hours are:

- Monday-Thursday 8am-midnight
- Friday 8am-8pm
- Saturday 9am-5p
- Sunday 8am-midnight
- Canvas technical requirements: https://clear.unt.edu/supported-technologies/canvas/requirements
- Other related hardware or software necessary for the course: such as headset/microphone for synchronous chats, word processor, etc.

Minimum Technical Skills Needed

Using the Internet and the learning management system Canvas, using email with attachments, creating and submitting files in commonly used word processing program formats, downloading and installing software, using MySQL programs.

Student Academic Support Services

- Code of Student Conduct: provides Code of Student Conduct along with other useful links
- Office of Disability Access: exists to prevent discrimination based on disability and to help students reach a higher level of independence
- Counseling and Testing Services: provides counseling services to the UNT community, as well as testing services; such as admissions testing, computer-based testing, career testing, and other tests
- UNT Libraries
- UNT Learning Center: provides a variety of services, including tutoring, to enhance the student academic experience
- UNT Writing Center: offers free writing tutoring to all UNT students, undergraduate and graduate, including online tutoring
- Succeed at UNT: information regarding how to be a successful student at UNT

ASSESSMENT & GRADING

Assessments

A student's grade is composed of the following:

- Attendance (5%)
- Assignments (40%)
- Quizzes (20%)
- Term Project (35%)

Grading

Class Attendance and Participation (5%). Students are strongly encouraged, but not required to attend online class meetings. Prior to the meeting, please preview the readings for the class and prepare your questions for discussion. The class will require students to post their thoughts/comments/answers to **at least 5** discussion questions. The questions will be given in each lesson. The posting should be in 50-100 words. It will be due in two weeks on Thursday midnight after the questions are posted.

Assignments (40%)
The class will have **FOUR assignments**. The assignments are designed to help students understand important concepts and gain hands-on experience in Python programming, data processing, and problem solving. Assignments should be typewritten, and diagrams should be drawn using graphics software packages such as PowerPoints and Excel.

Quizzes (20%)
There will be **TEN quizzes** for this course at the specified dates in Table 2. The quizzes will be made available to students on Canvas ONE day before that week's class. It will be due at 5:00 pm on the day of the class (Thursdays). Students need to complete the quiz in 30 minutes.

Term Project (35%)
(Report One 5%; Final Report and the System/programs: 20%; Presentation and evaluation 5%) The purpose of the term project is to apply what has been taught in this course to process real-world data or information. Students will work in teams (no more than 3 people) to tackle one particular problem assigned or approved by the instructor in Natural Language Processing, Information Visualization, Machine Learning, and/or Data Mining. Term project topics and their specifications will be distributed and discussed in class.

Total Points Possible for Semester/Grading Scale = 1000

1000-900 = A	*899-800 = B*
799-700 = C	*699-600 = D*
599 and below = F	

Grading Table

Assignment	Points Possible	Percentage of Final Grade
Attendance	*100 points*	*5%*
Assignment 1 –	*100 points*	*10%*
Assignment 2 –	*100 points*	*10%*
Assignment 3 –	*100 points*	*10%*
Assignment 4 –	*100 points*	*10%*
Quizzes • *10 quizzes @ 20 points ea.*	*200 points*	*20%*
Term Project • *Report one @ 50 points.* • *Final report @ 200 points.*	*300 points*	*35%*

• ***Presentation @ 50 points.***		
Total Points Possible	*1000 points*	*100%*

COURSE CALENDAR

The contents of the course are organized into 16 weeks. Please refer to Table 1 for lessons, topics, and readings materials. Table 2 lists the suggested study schedule, assignments, quiz, and term project due dates.

Table 1. Lessons and Readings

Lessons	Topics	Readings
Lesson 1	Orientation and overviews: Introduction to Python, iPython, the command line, Jupyter notebooks, Google Colab, GitHub.	Downey: Chapter 1
Lesson 2	Python Basics (1): Variables, Statements, Data Types, and Functions	Downey: Chapter 2-3, 8
Lesson 3	Python Basics (2): Conditionals, Recursion, and Advanced Functions	Downey: Chapter 5, 6, 7
Lesson 4	Python Basics (3): Iteration and File Operations	Downey: Chapter 7, 9, 14
Lesson 5	Python Basics (4): Data Structures: Lists, Tuples, Dictionaries, and Sets	Downey: Chapter 10-13
Lesson 6	Python Basics (5): Classes and Objects	Downey: Chapter 4, 15-18
Lesson 7	Python Basics (6): Exceptions, Debugging, and Computer Programming Tips	Downey: Chapters 20-21
Lesson 8	Python Web Crawling	Ryan: Chapter 1-5
Lesson 9	Python Database Operation	Ryan: Chapter 6
Lesson 10	Python Data Visualization	Igor: Chapter 1-3

Lesson Related Materials

- Lesson one:
  - ➢ Coding environment: Google Colab (http://colab.research.google.com/)
  - ➢ MOOC courses:
    - ✓ Python for Everybody Specialization on Coursera: https://www.coursera.org/specializations/python
    - ✓ Python for Data Science (PY0101EN) IBM: https://courses.cognitiveclass.ai/courses/course-v1:Cognitiveclass+PY0101EN+v2/info
    - ✓ Introduction to Python for Data Science, EdX, Microsoft and DataCamp: https://www.edx.org/course/introduction-python-data-science-3
    - ✓ Tutorials of Google Colab: https://www.youtube.com/watch?v=inN8seMm7UI
    - ✓ Success Story of Sylvain Gugger: https://www.fast.ai/2019/01/02/one-year-of-deep-learning/
  - ➢ To get help:
    - ✓ Python Docs: https://docs.python.org/3/
    - ✓ Python Forums: https://python-forum.io/
    - ✓ Stackoverflow: https://stackoverflow.com/

- Lesson eight:
  - ➢ Scrapy: A Fast and Powerful Scraping and Web Crawling Framework (http://docs.scrapy.org/en/1.6/index.html)
- Lesson nine:
  - ➢ SQL Solution of chapter23: https://colab.research.google.com/drive/1S_qUBL3cqex8A5sNw-GZ3eE-XvFOxIch
  - ➢ Python MySQL on W3Schools: https://www.w3schools.com/python/python_mysql_getstarted.asp
- Lesson ten:
  - ➢ Python Matplotlib: https://matplotlib.org/

Study Schedule and Due Dates
(Discussions will be due in two weeks on Thursday midnight after the questions are posted. Assignments and Project One will due on Thursday midnight of the specified week. Quizzes will be available online from 8:00 am on Wednesday to 5:00 pm on Thursday of specified week. Term project final report will due on December 13 midnight)

Table 2. Study Schedule and Due Dates

Academic Week	Dates	Meeting Date (Online)	Study Focus	Assignment/Project /Survey/Quiz Due
1	August 26-September 1	**August 29**	Syllabus, Lesson 1	Class Survey (not graded)
2	September 2-8	**September 5**	Lesson 2	**Quiz 1**
3	September 9-15	**September 12**	Lesson 3	**Quiz 2**
4	September 16-22	**September 19**	Lesson 4	**Assignment 1** **Quiz 3**
5	September 23-29	**September 26**	Lesson 5	**Quiz 4**
6	September 30-October 6	**October 3**	Lesson 6	**Assignment 2** **Quiz 5**
7	October 7 - 13	**October 10**	Lesson 7	**Quiz 6**
8	October 14 - 20	**No Meeting**	Lesson 1-7	**Assignment 3**
9	October 21- 27	**October 24**	Lesson 8	**Project Topics** **Quiz 7**
10	October 28 - November 3	**October 30**	Lesson 9	**Quiz 8**
11	November 4 - 10	**Meeting with your team**		**Term Project 1st Deliverable**
12	November 11- 17	**November 13**	Lesson 10	**Quiz 9**
13	November 18- 24	**No Meeting**	Review, Thanksgiving	**Assignment 4**
14	November 25- December 1	**Meeting with your team**	Work on Term Project	**Quiz 10**
15	December 2 – 8	**Meeting with your team**	Work on Team Project	
16	**December 13**	**December 11 Online meeting for all students**	Term Project Presentation	**Term Project Final Report Due at December 13 Midnight**

COURSE EVALUATION

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 at http://spot.unt.edu/ or email spot@unt.edu.

COURSE POLICIES

Assignment Policy

Students should submit the assignments and term project reports via Dropbox at class site in canvas.unt.edu: doc (or .docx) files and python files, also with the code uploaded on GitHub, details will be included in each assignment.

Examination Policy

There are no exams for this course.

Instructor Responsibilities and Feedback

- Helping students grow and learn
- Providing clear instructions for projects and assessments
- Answering questions about assignments
- Identifying additional resources as necessary
- Providing grading rubrics
- Reviewing and updating course content
- The instructor and TA will respond to students' emails and questions posted to the discussion boards within two days except for the weekends
- Assignments grades and feedbacks will be returned to the students within one week after the submission deadline.

Late Work and Missed Work

Students are expected to submit assignments and projects on time. **The due dates are Thursdays 11:59pm of the week specified in Table 2. Study Schedule and Due Dates**. If an extenuating circumstance such as a medically diagnosed illness or family emergency arises, which prevents you from submitting your assignments, you should contact the instructor and the TA as soon as possible before the due date. Late work without the permission of the instructor will receive a grade with a 10% penalty (or 10 points out of 100) per day after the due date. A student who is having trouble with assignments is strongly encouraged to contact the instructor and the TA as early as possible for personal advising.

Course Incomplete Grade

The UNT Graduate Catalog (http://catalog.unt.edu/index.php?catoid=16) describes and explains grading policies. A grade of Incomplete (I) will be given only for a justifiable reason and only if the student is passing the course. The student is responsible for meeting with the instructor to request an incomplete and discuss requirements for completing the course. If an incomplete is not removed within the time frame agreed to by instructor and student, the instructor may assign a grade of F.

Withdrawal

The UNT Graduate Catalog (http://catalog.unt.edu/index.php?catoid=16) describes and explains withdrawal policies and deadlines. The UNT semester course schedule lists specific deadlines regarding withdrawal. A grade of Withdraw (W) or Withdraw-Failing (WF) will be given depending on a student's attendance record and grade earned. Please note that a student who simply stops attending class and does not file a withdrawal form may receive an F.

Attendance Policy

Students are encouraged to attend each class meeting. Prior to the meeting, please preview the readings for the class and prepare your questions for discussion. You will miss class work and activities if you do not attend the class.

Students' Responsibility for Their Learning

The students are required to follow course schedule and finish the class work, assignments, quizzes, and term projects. Students are expected to study 12-15 hours per week to achieve satisfactory class performance. Students do not have programming experience are required to find extra materials to study.

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 at 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 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. 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/.

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.

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 at 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 through the use of 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 or her need to participate in an on-campus experiential component for this course, s/he 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.

Use of Student Work

A student owns the copyright for all work (e.g. software, photographs, reports, presentations, and email postings) he or she creates within a class and the University is not entitled to use any student work without the student's permission unless all of 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 of 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