Computational techniques for linguists[1]

Course Information

> This course aims to prepare students for careers in computational linguistics, where some basic software engineering skills are required.

Course objectives

In this course students will design and implement a project using industry best practices, with a focus on designs that can scale up to meet the needs of a large number of potential users. We will discuss topics such as test-driven development, object oriented programming, databases, web scraping, and RESTful APIs. There will also be a discussion of how to prepare for a technical job interview. By the end of the course, students will have a project that they can show to potential recruiters and have a plan to prepare for getting their first job.

Learning outcomes

By the end of this course, successful students will...

1. design and implement a complete NLP software application using industry best practices.[a,b,c]

2. be able to use and explain strategies for building software that can function at industrial scale.[a]

3. be well prepared for a technical job interview.

[a] relates to Linguistics HLT program outcome #1: *Students will demonstrate programming skills for the workplace.*

[b] relates to Linguistics HLT program outcome #2: *Students will be able to use fundamental algorithms and concepts in Natural Language Processing.*

[c] relates to Linguistics HLT program outcome #3: *Students will show knowledge of tools and packages used in Natural Language Processing.*

Prerequisites

Some familiarity with programming in Python is expected, such as basic flow control, looping, and built-in data structures such as lists and dictionaries. We will be using Python 3.9 in this course.

---

[1] The original development of this course was done by Jeffrey Berry.

Instructor

name	Eric Jackson
email	ejackson1@email.arizona.edu
office hours	Thursdays 10:00am–12:00pm (Arizona time, UTC-7) and by appointment, online via Zoom at `https://arizona.zoom.us/j/85345960433` (passcode 093770)

This is an asynchronous course, so we won't meet at a single time for class to happen. There are no sessions at which I will take attendance. My working hours are normal business hours in Arizona, and I generally do my preparation and grading for the course at that time. If you need to contact me, the best way is to send me an email or a forum message. I try to respond within 24 hours, but *during* working hours, I may be grading coursework, meeting with someone, or recording lectures, so my response may not come immediately. If there's a chance you may need my response, don't wait to discover that when a deadline is upon you.

I know that normal business hours in Arizona are not convenient for everyone in this course. If you need to meet outside of Arizona business hours, I'm holding Thursday evenings open. Contact me in advance to set up a time and a link; I won't otherwise be online then.

Course forum (Zulip)

For this course, we'll use a forum that is outside the course website, though still provided by the university. Make sure you sign up in the first week. Please be respectful of others in your interactions on the forum.

`https://forum.hlt.arizona.edu/#narrow/stream/27-ling-508-su2022`

In addition to the already-planned course forum activities, you're free to start discussions there with the class, or post questions about the course. Bear in mind that responses from the forum (from me or from other students) may come quickly, but it is not guaranteed that you will receive responses faster than one working day. You should plan for deadlines as if forum responses will take 24 hours or more.

Readings

There is one required textbook that **you** are responsible to obtain a copy of.

Laakman McDowell, Gayle. 2015. *Cracking the Coding Interview*, 6th edn. CareerCup.

ISBN-13: 978-0984782857

There will be other supplemental readings each week, assigned through the course website. All other readings are available online for free (digitally) through the UA library. You'll need to log in to UA Libraries or the O'Reilly Learning site with your NetID and password to access these other readings.

Course schedule

The course is divided into seven topical modules, one each week, with lectures, readings, and assignments for each module. All course material are available on D2L or are linked from within D2L. Specific pages for each official reading assignment will also be found in D2L. All dates and times for the course are in Arizona Mountain Standard Time (MST-7).

START DATE	UNIT TOPIC
7/5/22	Unit 1: Introduction; Test-driven development; Scalability
7/11/22	Unit 2: Object-oriented programming; Universal Modeling Language (UML)
7/18/22	Unit 3: Web scraping
7/25/22	Unit 4: Databases & SQL
8/1/22	Unit 5: Services & orchestration
8/8/22	Unit 6: RESTful API; User interface
8/15/22	Unit 7: Job interviews

The first day of the course is July 5, 2022, and the last day of the course is August 19, 2022. All work must be submitted by the last day of the course.

Requirements and grading

Students are expected to actively participate in the course by watching the recorded lecture videos, reading and digesting the assigned readings, completing the assigned work, and engaging with the instructor and other students in the course forum. You are all adults, and I expect you to take responsibility for your own learning.

Technology

This is a fully online course. As such, you will need a stable internet connection to access course content and submit assignments. To complete your assignments, we recommend that you use a computer with >6GB of RAM. To complete your assignments, you will also need

- Git

- A GitHub account

- Docker

- A modern web browser (Firefox or Chrome recommended)

- PyCharm https://www.jetbrains.com/pycharm/ (You can use the community edition or get a student license)

For students who need help with R, Python, GIS, Data Management and Data Visualization, the University of Arizona Library's iSpace provides free virtual drop-in hours throughout the semester. Times vary by semester; if you're interested, check the library website to follow up. `https://libcal.library.arizona.edu/calendar/events/data-viz-drop-in?hs=a`

ResBaz Arizona hosts a variety of events providing researchers opportunities to connect with experienced data scientists and engineers. For more information, see their homepage: `https://researchbazaar.arizona.edu/#portfolio`

Required reading

Readings are assigned for each unit of the course. You may read the assigned sections before you watch the lectures, or after—it's up to you. However, you must read the assigned sections and complete any activities before the end of that week.

Videos and lectures

Lecture videos will be provided through Panopto on the course website (D2L). Some required videos are hosted on sites outside D2L, such as at Youtube.com; in such cases, links to external sites will be provided within D2L. You are expected to watch all videos and understand the content. If you for any reason cannot access a required video, please notify the instructor immediately.

If the content of a unit is not clear to you on the first viewing, don't panic. Make sure you've done the required readings and try re-watching the lectures. You're free to search for other presentations of the same topic online. If a concept is still unclear, you are expected to send a question to the instructor by email, meet with the instructor in regular office hours or arrange another time to meet, or post a question for clarification on the course forum.

Assignments and grading

There are 360 points available in the course: 340 are required and 20 are bonus points.

There will be one graded quiz each week, available via D2L (quizes are 10 questions, 1 point per correct answer, for a total of 70 points for the course; quizes can be retaken 10 times).

There will be six programming assignments, each worth 20 points. The programming assignment in Week 3 will be a bonus assignment. Programming assignments will be submitted and graded through GitHub. The full 20 points will be given for submissions that pass all tests (ie, are given a green check in GitHub), and 10 points will be given for submissions that do not pass the tests (ie, are given a red X in GitHub).

There will be one course-long project in which you will apply what you learn in this course to design a web-based application from the ground up. There will be graded milestones that are due each week, and you will also create a demonstration video for your application. At the end of the course, you'll have a project for your portfolio or that you can show to potential employers. Altogether this project will constitute 170 points, or half your grade for the semester. Detailed descriptions of the point break-down for each week are given in the Project module in D2L

- Week 1: Use Cases Document (10pts)
- Week 2: UML Class Diagrams (10pts)

- Week 3: Code uses a test and OOP (20pts)

- Week 4: Code uses a database (20pts)

- Week 5: Code has at least 1 function that implements a use case from #1 (20pts)

- Week 6: Code has a Flask API server (20pts)

- Week 6: Documentation Page (Week 6) 20 pts

- Week 7: Code is complete and has a simple HTTP form page (20pts)

- End of course: Demo Video (30pts)

All your quizes and assignments are graded on a point basis, with the number of points available specified for each assignment. Grades will be curved, and a final grade of A, B, C, D, or E will be given. The following minimum percentages will guarantee the corresponding letter grades, in accordance with university policy:

A: 90–100%
B: 80–89.9%
C: 70–79.9%
D: 60–69.9%
E: 0–59.9%

Grades will be accessible within D2L. More details about the programming assignments and project are posted on D2L. Based on these point values, the quizes, assignments, and project will contribute to your final grade as follows:

type	number	point total (percent of total)
D2L quizes	7	70 (21%)
programming assignments	5 + 1 bonus	100 + 20 bonus (29% w/o bonus)
project	1	170 (50%)
total		340 + 20 (100% up to 105.9%)

The due date for each assignment will be posted with the assignment in D2L. All times will be given in Arizona time (Mountain Standard, GMT-7). If you have an unexpected life event that will keep you from completing an assignment on time, talk to me about accomodation as soon as you can. **Late work will otherwise not be accepted.**

Collaboration Policy

Students are encouraged to discuss problems and general approaches for solutions, but everyone must turn in their own work. Using some code you found online is acceptable, but cite any sources you use, and make sure that the majority of the work you turn in is your own. You may not submit assignments that are substantially the same as your classmates. Do your own work, and please ask me in advance if you are unsure whether something will be acceptable or not.

Covid

1. The university has a specific site for covid information: `http://covid19.arizona.edu`.

2. I understand that although we're all ready to get back to normal, these are still extraordinary times, and you may still be experiencing new personal and financial challenges. Let me know if we need to make accommodations for covid-related things.

University boilerplate

Disabilities If you have a disability that affects how you will need to do the work in this class, please let me know *within the first week of class.*

Academic Code of Conduct Cheating and plagiarism are not remotely acceptable in any way. By this point in your academic career, it is important that you know what these things are and know what you are and are not permitted to do. You can find more information in the UA Code of Academic Integrity. Disruptive or disrespectful behavior on the class website or forum is not acceptable.

Sensitive Material This is a university and you are adults. It is possible that we may touch on topics that some students could find sensitive during the semester. Given the focus of this course, this seems unlikely, but I alert you nonetheless.

University policies

All of the following are things the university requires us to put on syllabi.

Absence and Class Participation Policy

Attendance for asynchronous courses works differently than it does for synchronous courses, but the university requires the following on syllabi anyway.

The UA's policy concerning Class Attendance, Participation, and Administrative Drops is available at: `http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop`

The UA policy regarding absences for any sincerely held religious belief, observance or practice will be accommodated where reasonable, `http://policy.arizona.edu/human-resources/religious-accommodation-policy`.

Absences pre-approved by the UA Dean of Students (or Dean Designee) will be honored. See: `https://deanofstudents.arizona.edu/absences`

Classroom Behavior Policy

To foster a positive learning environment, students and instructors have a shared responsibility. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. To that end, our focus is on the tasks at hand and not on extraneous activities (e.g., texting, chatting, reading a newspaper, making phone calls, web surfing, etc.).

Students are asked to refrain from disruptive conversations with people sitting around them during lecture. Students observed engaging in disruptive activity will be asked to cease this behavior. Those who continue to disrupt the class will be asked to leave lecture or discussion and may be reported to the Dean of Students.

Threatening Behavior Policy

Required language: The UA Threatening Behavior by Students Policy prohibits threats of physical harm to any member of the University community, including to oneself. See `http://policy.arizona.edu/`
`education-and-student-affairs/threatening-behavior-students`.

Accessibility and Accommodations

Recommended language is provided on the Disability Resource Center website: `http://drc.arizona.edu/instructors/syllabus-statement`.

Code of Academic Integrity

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See: `http://deanofstudents.arizona.edu/academic-integrity/students/academic-integrity`.

UA Nondiscrimination and Anti-harassment Policy

The University is committed to creating and maintaining an environment free of discrimination; see

`http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy`

Subject to Change Statement

Information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructor.