CIS 5300: Natural Language Processing
Summer 2024

Instructors

Dr. Chris Callison-Burch, Department of Computer and Information Science
Dr. Kuzman Ganchev, Lecturer

Course Description

This course provides an overview of the field of natural language processing. The goal of the field is to build technologies that will allow machines to understand human languages. Applications include machine translation, automatic summarization, question answering systems, and dialog systems. NLP is used in technologies like Amazon Alexa and Google Translate.

Course Prerequisites

Required: CIT 5910, 5920, and 5940
Suggested: CIT 5960 (students can take as a co-requisite)

Course Textbook

Speech and Language Processing (3rd ed. draft), by Dan Jurafsky and James H. Martin
- Available for free from [the authors' website](#)

Grading & Assessments

You must attempt all graded assignments to pass the course. If you have any questions or concerns about grading or progress in the course, please reach out to your instructor.

This course will use a variety of assessments to determine whether learners understand and can apply the key concepts and skills that the course teaches. This includes:

Type	%	Description
Homework assignments	75%	There are nine individual assignments or "Homeworks," in this course. These are the main assessments in this course.
Timed Exam	15%	This course contains one timed, proctored exam.
Quizzes	10%	This course contains one quiz per module.
Extra Credit	<= 3%	We periodically offer extra credit assignments related to AI research. These are entirely optional and cannot negatively impact your grade. Because the class is not curved, even if your classmates do the extra credit and you do not, your grade will not be negatively impacted. Extra credit is capped at a maximum of 3% towards your final grade.

Grading/Curving

This course is not curved. Your overall score is computed as *0.15* timed exam score + 0.75 * total homework score + 0.1 * total quiz score + optional extra credit*. Here is how letter grades are assigned based on your overall score:

Score	Grade
>= 98	A+
93-98	A
90-93	A-
87-90	B+
83-87	B

80-83   B-
75-80   C+
70-75   C
65-70   C-
50-65   D
below 50 F

Please be aware that although you need >= 98 in order to get an A+, we will only be awarding an A+ to - at most - 10% of the class.

Late Policy

If students submit their assignments later than the stated deadline, they will face a 10% penalty per late day, till their score hits zero. For extraordinarily extenuating circumstances, please fill out the extension request form located under the Course Resources page in Canvas. Please fill out this form for any and all extension requests you wish to submit. These extension requests must be submitted at least 24 hours before the assignment deadline. **If your request is granted, you will see updated deadlines reflected in Canvas and Gradescope by Mondays at 5 pm ET.** You will need to fill out this form for each assignment you wish to have an extension for.

Regrade Requests

Regrade requests are allowed up to 1 week after the grades are released. These requests must be made through Gradescope.

Extra Credit

Any extra credit opportunities will be announced during the course on Ed Discussion.

Recitation

This course will contain recitations. Recitations will be announced on EdDiscussion.

Getting and Giving Help

TA and Faculty Support

TAs will hold office hours weekly where they will open a queue in PennLab's OHQ.io system. Please note that students should not employ TAs as private tutors. Your professor will hold

open office hours on a regular basis. The TA and faculty sessions will be listed in the Canvas calendar.

Collaboration Guidelines

In the professional world of software development, collaboration—including using code that others have written—is both practical and ubiquitous.

However, to prepare for entering that professional context, you need to develop a full set of software development skills so that you are both able to create your own code and evaluate the quality of someone else's code that you might use.

In the context of this course, independent work and evaluation are critical. **Do not collaborate with others on individual graded assignments unless it is explicitly indicated.** The inappropriate collaboration will be considered cheating and considered under Penn's Code of Academic Integrity.

Discussion forums *are* collaborative—please take advantage of those times to work with your colleagues. For general communication with your colleagues, use your Slack channels or Slack direct messages.

Forming study groups to understand the material is also a good idea, as long as you stay on the conceptual level and are *not* collaborating on the graded assignments directly.

Note: When in doubt always ask the instructor or TA first, to avoid any potential collaboration that can lead to academic dishonesty.

Do not cheat. **Please note that searching for solutions online is the same as cheating**. **Copying code is considered cheating. Posting solutions online is also considered cheating.** If you are caught posting solutions or code to a **publicly accessible location** (like StackOverflow or GitHub), it will be considered cheating. If you do use GitHub (or similar cloud-based code management system) to set up a remote code repository, **YOU ARE REQUIRED TO KEEP THAT REPOSITORY PRIVATE**.

Downloading course materials is only for your own academic use. You are not permitted to distribute these materials (including posting them online.)

You can further read Penn's Code of Academic Integrity page on this subject matter, as well as the SEAS Graduate Student guidelines on the code of ethics.

Plagiarism Policy

The first instance of homework plagiarism will be handled by the instructor and may include escalation to the Center for Community Standards and Accountability (CSA).

Second instances or exam plagiarism will be turned over immediately to the University of Pennsylvania's Center for Community Standards and Accountability.

Regardless of previous work in the course, the penalty for plagiarism is the failure of the course (regardless of current course average), and potential permanent notation on your academic record that will follow you to all future academic institutions and possibly future employers. If you are unfamiliar with what constitutes plagiarism at Penn, visit Penn's Code of Academic Integrity.

Access to Materials and Content Before and After Graduation

Access to course materials and your submissions is not guaranteed after the completion of a course. Therefore, we recommend that students download any assignments or materials they would like to keep before a course concludes.

Recording Notice

Public office hours, recitations, and other live events will be recorded, used, and may be made available to class participants during the current semester as well as students who take the class in future semesters.

Private office hours will also be offered and are not recorded. Students who do not wish to attend the publicly-recorded office hour may attend the private office hours.

Summer 2024 Course Schedule and Important Dates

Dates are subject to change. Please check EdDiscussion for announcements regarding schedule changes. Note that weeks run Monday through Sunday.

	Course #5300 | Summer 2024						
	Note: Weeks run Monday through Sunday						
	SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
Week 1	5/5	5/6	5/7	5/8	5/9	5/10	5/11

		First day of term					
Week 2	5/12	5/13	5/14	5/15	5/16	5/17	5/18
		Add Deadline					
		[1.0 CU Courses and 0.5 CU 1st Seven Weeks]					
Week 3	5/19	5/20	5/21	5/22	5/23	5/24	5/25
		Spring 2024 grads- official graduation date					
		Homework 1 Due					
Week 4	5/26	5/27	5/28	5/29	5/30	5/31	6/1
		Memorial Day (No Classes)	Homework 2 Due			Course Drop Deadline	
						[1.0 CU Courses and 0.5 CU 1st Seven Weeks]	
Week 5	6/2	6/3	6/4	6/5	6/6	6/7	6/8
		Homework 3 Due					

Week 6	6/9	6/10	6/11	6/12	6/13	6/14	6/15	
		Homework 4 Due				Late Course Drop Deadline		
						[1.0 CU Courses and 0.5 CU 1st Seven Weeks]		
Week 7	6/16	6/17	6/18	6/19	6/20	6/21	6/22	
		Homework 5 Due		Juneteenth (No Classes)		Last Day to Withdraw from a Course		
						[0.5 CU 1st Seven Weeks]		
Week 8	6/23	6/24	6/25	6/26	6/27	6/28	6/29	
	Last Day of Classes	First Day of Classes						
	[0.5 CU 1st Seven Weeks]	[0.5 CU 2nd Seven Weeks]						
Week 9	6/30	7/1	7/2	7/3	7/4	7/5	7/6	
		Add Deadline			Independence Day (No Classes)			
		[0.5 CU 2nd Seven Weeks]						
		Homework 6 Due						
Week 10	7/7	7/8	7/9	7/10	7/11	7/12	7/13	

Week 11	7/14	7/15	7/16	7/17	7/18	7/19	7/20
		Homework 7 Due					
Week 12	7/21	7/22	7/23	7/24	7/25	7/26	7/27
				Fall 2024 Registration Opens			
Week 13	7/28	7/29	7/30	7/31	8/1	8/2	8/3
		Homework 8 Due					
Week 14	8/4	8/5	8/6	8/7	8/8	8/9	8/10
		Homework 9 Due				Summer 2024 grads- Official graduation date	
Week 15	8/11	8/12	8/13	8/14	8/15	8/16	8/17
	Last day of term						