Ling 317 Introduction to Computational Linguistics
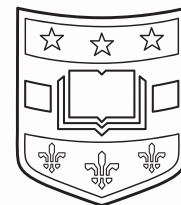
Dr. Nick Danis, nsdanis@wustl.edu
Washington University in St. Louis
https://www.nickdanis.com/syllabi/sp2021-ling317

Course Info

Course Number	L44 Ling 317
Semester	Spring 2021
Time	Tu/Th 1-2:15pm
Location	Zoom*

*all Zoom links available on Canvas

Description

This course introduces the computational tools, both practical and theoretical, to describe and analyze natural language. We will touch upon most of the major aspects of the field of computational linguistics and natural language processing, such as analysis of raw corpora; modeling phonological, morphological, & syntactic patterns; and learnability. We will use Python 3 throughout the course. This course is an introduction to these topics, and therefore no programming experience is required. (In fact, this course is *not* meant for students with a strong background in computer science.) Students, however, are encouraged to complete optional exercises and readings on basic programming techniques to aid in the completion of assignments and discussions. Prerequisite: L44 Ling 170D and either L44 Ling 258 or CSE 131.

Course Goals

1. Learn practical, computational tools (e.g. Python NLTK) to manipulate natural language data
2. Understand the computational properties of natural language, independent of whether it is being computed by machines or humans (i.e. treating grammars as purely mathematical objects)
3. Apply these techniques to solve problems in phonology, morphology, and syntax

Required Materials

There are no required purchases for this course. Any readings will be made available online. Additionally, all software necessary is free as well. This course assumes you have (or have access to) a working computer with a desktop operating system (macOS, Windows, Linux, etc.). For those

students working primarily off a Chromebook or iOS/Android device, there are cloud options available. More details are given on Canvas.

Readings

A majority of the readings will come from the following source:

- Steven Bird et al. (2014). *Natural Language Processing with Python*. 2nd ed. URL: http://www.nltk.org/book/

The entire second edition is available online at the link above; do not purchase a physical copy as the printed edition is the outdated first edition.

Software

All course material that is not a PDF is distributed in a Jupyter notebook for Python 3. If you are unsure what to install, or what this means, we will go over it the first week of class. There is detailed instructions on the Canvas site as well. If you have some experience, and have Python 3 installed, you can get a headstart with `pip install jupyter`.

Special covid-19 related policies

This class is remote-only, synchronous with an asynchronous option. On the schedule below, topics appear in pairs: the first day is a lecture, and the second day is hands-on coding. If you are asynchronous, I would recommend trying to attending the coding sessions live as much as possible, though I understand the difficulties in doing so especially if there is a substantial timezone difference. If at any point you feel you are having any sort of difficulties with the course, please do not hesitate to reach out.

Grade

The grade breakdown is shown below.

Category	Weight
Participation & Discussion	15%
Weekly Assignments*	60%
Final Project	25%

*The lowest grade is dropped.

Participation entails completing & uploading the in-class notebooks to Canvas once completed, and regular engagement on the course Piazza & in-class discussions.

Letter grades

Letter grades are assigned based off the following scale. Numerical grades are **not** rounded.

$$100 \geq A+ \geq 98 \qquad\qquad 80 > C+ \geq 77$$
$$98 > A \geq 93 \qquad\qquad 77 > C \geq 73$$

$$93 > A- \geq 90 \qquad 73 > C- \geq 70$$
$$90 > B+ \geq 87 \qquad 70 > D+ \geq 67$$
$$87 > B \geq 83 \qquad 67 > D \geq 63$$
$$83 > B- \geq 80 \qquad 63 > D- \geq 60$$

If you are taking this class pass/fail, you must receive at least a C- (70%) to pass.

If you believe there has been an error in grading, I am happy to discuss it with you. However, you must bring it up to me within one week of the graded assignment being returned to you. After this, the grade is considered final.

Schedule

The exact schedule might change as the class progresses. Please see Canvas for all readings and assignment due dates.

Date	Day	Week	Comment	Unit
1/26/2021	Tuesday	1		Introduction
1/28/2021	Thursday	1		Introduction
2/2/2021	Tuesday	2		Chatbots
2/4/2021	Thursday	2		Chatbots
2/9/2021	Tuesday	3		Regular expressions
2/11/2021	Thursday	3		Regular expressions
2/16/2021	Tuesday	4		Text Processing
2/18/2021	Thursday	4		Text Processing
2/23/2021	Tuesday	5		n-grams
2/25/2021	Thursday	5		n-grams
3/2/2021	Tuesday	6	Wellness Day	**No Class**
3/4/2021	Thursday	6		tf-idf
3/9/2021	Tuesday	7		Language Modeling
3/11/2021	Thursday	7		Language Modeling
3/16/2021	Tuesday	8		Finite State Machines
3/18/2021	Thursday	8		Finite State Machines
3/23/2021	Tuesday	9	Study Day	**No Class**
3/25/2021	Thursday	9		Finite State Machines
3/30/2021	Tuesday	10		Finite State Transducers
4/1/2021	Thursday	10		Finite State Transducers
4/6/2021	Tuesday	11		Context-free Grammars

Date	Day	Week	Comment	Unit
4/8/2021	Thursday	11		Context-free Grammars
4/13/2021	Tuesday	12		Pushdown Automata
4/15/2021	Thursday	12		Pushdown Automata
4/20/2021	Tuesday	13		Presentations
4/22/2021	Thursday	13		Presentations
4/27/2021	Tuesday	14	Study Day	**No Class**
4/29/2021	Thursday	14		Presentations
5/4/2021	Tuesday	15	Last Day of Classes	Presentations

Academic Integrity

This course adheres to the university's Academic Integrity Policy (https://studentconduct.wustl.edu/academic-integrity), and takes cheating and plagiarism very seriously. All work completed online must be done alone unless instructed otherwise, and no resources not approved by the instructor may be used during exams.

ADA Compliance

Washington University is committed to providing accommodations and/or services to students with documented disabilities. Students who are seeking support for a disability or a suspected disability should contact Disability Resources at 935-4153. Disability Resources is responsible for approving all disability-related accommodations for WU students, and students are responsible for providing faculty members with formal documentation of their approved accommodations at least two weeks prior to using those accommodations. I will accept Disability Resources VISA forms by email and personal delivery. If you have already been approved for accommodations, I request that you provide me with a copy of your VISA within the first two weeks of the semester. Please see more information at http://cornerstone.wustl.edu.

Sexual Assault Resources

The University is committed to offering reasonable academic accommodations (e.g., no contact order, course changes) to students who are victims of relationship or sexual violence, regardless of whether they seek criminal or disciplinary action. If you need to request such accommodations, please contact the Relationship and Sexual Violence Prevention Center (RSVP) at rsvpcenter@wustl.edu or 314-935-3445 to schedule an appointment with an RSVP confidential, licensed counselor. Information shared with counselors is confidential. However, requests for accommodations will be coordinated with the appropriate University administrators and faculty. Please see more information at https://students.wustl.edu/relationship-sexual-violence-prevention-center.