

Computational Methods in Linguistics

Ling 484 – Spring 2021

Not the final version yet

Time	MTT 456	Course Website	Moodle
Place	Zoom		
Instructor	Ümit Atlamaz	TA	Utku Türk
Email	umit.atlamaz@boun.edu.tr	Email	utku.turk@boun.edu.tr
Office Hours	TBD	PS Hours	TBD

Course Description

This course is an introduction to the field of computational linguistics. Throughout the semester, we will analyze human language as a computational device, discuss some of the major theoretical aspects of computational linguistics, and implement practical applications of these theories in Python. The intended audience for this course is students with some background in linguistics but not much experience in computational methods or programming. This course is not intended for students with a strong background in computer science. Feel free to contact the instructor if you still want to take the course or audit some of the classes.

Course Learning Goals

- Analyzing natural language from a computational perspective
- Understanding some of the major symbolic and non-symbolic approaches in computational linguistics
- Applying some computational technology (i.e. nlp techniques) to natural language
- Learning some Python to do basic text processing

Programming

In this course, we will use Python 3 as the programming language. If you have never written a line of Python before, *great!* This course is just right for you. If you have some experience in Python, you might still benefit from practicing your skills on analyzing natural language via Python. If you are a Python Ninja, you might want to volunteer to be a mentor for your friends.

You will need a Python 3 interpreter for this course. I will use Google Colab, which requires no installation and works on any machine (mac, pc, linux, even ios and android I believe.). All you need is a Google account. If you'd like to use some other interpreter, feel free to use Jupyter Notebooks or whichever IDE you like. Please note that my support for anything other than Google Colab will be minimal and you'll need to let me know in advance if you have any issues or questions.

Reading

Throughout the semester, we will use chapters from the following books and some others. Relevant chapters will be posted on Moodle every week:

- Bird, Steven, Ewan Klein, & Edward Loper. *Natural Language Processing with Python*. URL: <https://www.nltk.org/book>
- Jurafsky, Daniel & James H. Martin. 2020. *Speech and Language Processing 3rd Edition*. URL: <https://web.stanford.edu/~jurafsky/slp3>

Course Requirements & Evaluation

Final grades will be determined by your performance on assignments, in class presentations, a final project, and your participation. Covid-19 has altered our lives significantly. All of us are being deeply impacted. Regardless, we need to work together to minimize its impact on our online class experience. For this reason, participation will be a key evaluation metric in your course grade and it is your responsibility to make sure you are participating enough.

Evaluation Criteria		Grade Catalog	
Assignments	40%	AA	≥ 90%
Proposal Presentation	10%	BA	≥ 85%
Participation	10%	BB	≥ 80%
Project Presentation	10%	CB	≥ 75%
Final Project	30%	CC	≥ 70%
		DC	≥ 60%
		DD	≥ 55%
		F	≤ 54%

Final Project

You have two options for your final project.

- A 5-page essay on a topic (theoretical or practical) discussed in the course. (1.5 spaced)
- A python project on a topic covered in the course submitted as a Colab or Jupyter Notebook.

Policies

Attendance

Attendance is mandatory. Any student who misses five or more classes is considered to have “missed more than an occasional class”. Students who must, for any reason, miss more than an occasional class should consult with their instructors directly.

Late Assignments & Projects

Late assignments and projects will not be accepted, except:

- If you are ill and missed the deadline. You will be asked to provide proof of illness (doctor’s report).
- If you know you will miss a deadline due to a family, cultural, or religious observance, it is your responsibility to contact your instructor early in the semester. You should arrange to take an alternative assessment, or to submit the assignment on another date.

Academic Integrity

Students are expected to know and follow Boğaziçi University policies on Academic Integrity:

- Student Rights and Responsibilities

Students may not collaborate on answering questions in homework assignments or exams; homework assignments and exams must be done independently. Collaboration is a violation of the Academic Integrity policy. All instances of plagiarism will be reported.

Disability Services


Students with disabilities requesting accommodations should reach out to the instructor as well as:

- Students with Disabilities Unit
- Engelli Öğrenciler Danışma Ve Koordinasyon Birimi

Please, make sure your email registered in ÖBİKAS is in use. It is your responsibility to keep track of the email announcements.

Tentative Schedule

Disclaimer: The schedule below is tentative and subject to change. We will shift gears depending on the progress we make. Some of you have expressed interest in topics like Sentiment Analysis, Parsing, Embeddings, Markov Models, etc. We will discuss this during the semester and adjust the syllabus accordingly. We might be able to squeeze in some of these topics as we make progress during the semester. Alternatively, we might be able to replace some of the topics upon popular demand. If you are deeply interested in one of these topics, you can choose to work with me to create a lesson on the topic and present it as your term project.

Date	Topic	Reading	Assignment
Week 1 March 22 - 23	Intro, Syllabus, Marr's Three Levels Python Setup	Marr 1982	
Week 2 March 29 - 30	Formal Languages, Grammars, Regular Expressions		
Week 3 April 5 - 6	Regular Expressions NLP Pipelines & Text pre-processing		Assignment 1
Week 4 April 12 - 13	Edit Distance Context Free Grammars (Generation)		
Week 5 April 19 - 20	Context Free Grammars (Parsing) N-grams		Project Proposal due
Week 6 April 26 - 27	N-grams		Assignment 2
Week 7 May 3 - 4	Finite State Automata		
Week 8 May 10 - 11	SPRING BREAK		
Week 9 May 18 - 18	Finite State Transducers		Assignment 3
Week 10 May 24 - 25	BOW, TF-IDF		
Week 11 May 31 - June 1	Naive Bayes & Document Classification		
Week 12 June 7 - 8	Naive Bayes & Document Classification		Assignment 4
Week 13 June 14 - 15	Project Presentations		

All assignments are due Saturdays by 5:00 pm.