

Syllabus

Note: The latest course syllabus can be found on the course webpage <https://nlp.cs.gmu.edu/course/cs499-spring22/syllabus/>

CS 499

Natural Language Processing (NLP)

Instructor

[Ziyu Yao](mailto:ziyuyao@gmu.edu) (ziyuyao [at] gmU [dot] edu)

Office Hours: by appointments (Virtual or in person at ENGR4415).

Teaching Assistant

TBD

Office Hours: TBD

Meets

Monday and Wednesday, 3:00 to 4:15 PM, Planetary Hall 206.

Safe Return to Campus: Students are expected to follow the university's [Safe-Return-to-Campus Policy](#) (including mask wearing, daily health check, etc.) for attending any classes. Please check out the policy before coming to the campus and the classroom. Note that students who choose not to abide by these expectations will be referred to the Office of Student Conduct for failure to comply.

Course Web Page

<https://nlp.cs.gmu.edu/course/cs499-spring22/>.

We will use **Blackboard** for course materials/assignments/grading, and **Piazza** for Q&A (sign up link: [TBD](#)).

Course Description

Massive amounts of information in our daily life are expressed in natural language. In this class, we will study building computing systems that can process, understand, and communicate in natural language. This field is called natural language processing, or NLP. This class will focus on introducing fundamental concepts in NLP, and will cover techniques and necessary programming skills for building machine learning/deep learning-based NLP models. In the last several classes, we will further study cutting-edge research problems in NLP, including text generation, question answering, neural network interpretation, interactive learning, multilingual NLP, and so on.

Prerequisites

CS310 (Data Structure), CS330 (Formal Methods and Models), and proficiency in Python programming. Please contact the instructor if you have questions about the necessary background.

Class Format

The class will be in-person. Each class will take the following format:

- *Reading:* Before the class, you will be pointed to some reading materials (see "Reading Materials" in the course schedule) . Reading is not required but highly recommended to do before the class.
- *Summary/Elaboration/Q&A:* In the class, the instructor will summarize important points from the reading material, elaborating on details that were not included in the reading while fielding any questions. New material on cutting-edge methods, or a deep look into one salient method will also be covered.
- *In-class Coding Exercise/Quiz:* In some classes, the instructor will provide coding templates, and the students will need complete coding exercises in class. Sometimes this will be a quiz with a couple of questions. The purpose of the exercises/quizzes is to help students get familiar with concepts covered in class and/or kick start their assignments. Student performance will be graded; therefore, class attendance is required.
- *Presentation:* In the last class, students need present their final projects in class.

Grading

There will be no midterm or final exam. Your final grade will be dependent on:

- In-class Coding Exercise/Quiz: 20%.
- Assignments: 40%. There will be four assignments in total. Each assignment will be an independent "small" coding project. In the project, the instructor will provide a code template and students will be instructed to complete the project.
- Final Project and Presentation: 40%. This includes (1) submitting a project proposal in the middle of the semester; (2) extending from your proposal, submitting an intermediate project report describing the baseline designs and experimental results, along with the source code implementation; (3) submitting the final project report with source code in the end of the semester; and (4) presenting your project in the last class.

See details about [the assignments](#) and [the final project](#) following the links.

Your final letter grade will be given based on (depending on class performance, the instructors may shift these boundaries down to raise students' grades.):

Letter Grade	Points (out of 100)
A	97-100
A-	90-96
B+	86-89

B	83-85
B-	80-82
C+	76-79
C	73-75
C-	70-72
D	60-69
F	0-59

Late Day Policy for Assignments:

In case there are unforeseen circumstances that don't let you turn in your assignment on time, 4 late days *total* over the four assignments will be allowed. Notes: (1) Late days cannot be applied to the final project; (2) The last two assignments are harder than the others, so it'd be a good idea to try to save your late days for them if possible; (3) The late days cannot be used fractionally, e.g., submitting the assignment 1 hour late will incur 1 late day. Assignments that are late beyond the allowed late days will be graded down by 5% per day. In the case of a serious illness or other excused absence, as defined by university policies (including providing necessary evidence), coursework submissions will be accepted late by the same number of days as the excused absence.

Class Attendance Policy:

As we will have coding exercises/quizzes for some classes, attendance is required. However, in the case of a serious illness or other excused absence, as defined by university policies (including providing necessary evidence), students will be excused and the exercise/quiz grade will be dropped. I expect such cases to be relatively rare, and if you'll be away for more than 2 classes over the semester, please consult in advance.

Readings

Students should be able to understand the course content just by following the lecture and by doing the readings. However, the following textbooks serve as good references.

- Jurafsky and Martin, Speech and Language Processing, 3rd edition [\[online\]](#) (Referred to as "JM");
- Jacob Eisenstein, Natural Language Processing [\[online\]](#) (Referred to as "Eisenstein").

Tentative Schedule

#	Date	Topic	Reading Materials	Assignments
1	01/24	Introduction and Class Outline		
2	01/26	Working with Text in Python		
3	01/31	N-gram Language Models		Assignment 1 Out
4	02/02	Classification 1		
5	02/07	Classification 2		
6	02/09	Classification 3		
7	02/14	Neural 1: Feedforward Neural Networks		Assignment 1 Due Assignment 2 Out
8	02/16	Neural 2: Word Embeddings, Neural Network Implementation		
9	02/21	Neural 3: Neural Language Models		
10	02/23	Sequence 1: Tagging, HMMs		
11	02/28	Sequence 2: Viterbi and Beam Search, CRFs		Assignment 2 Due Assignment 3 Out
12	03/02	Parsing 1: Dependency Parsing		
13	03/07	Parsing 2: Constituency Parsing		
14	03/09	Parsing 3: Semantic Parsing		
15	03/14	Spring Recess - No Class		Assignment 3 Due
16	03/16	Spring Recess - No Class		
17	03/21	Neural 4: Sequence-to-Sequence Models		Project Proposal Due Assignment 4 Out
18	03/23	Machine Translation		
19	03/28	Neural 5: Attention, Transformers		
20	03/30	Contextual Representations and Pre-training 1		
21	04/04	Contextual Representations and Pre-training 2		Assignment 4 Due
22	04/06	Text Generation 1		
23	04/11	Text Generation 2		
24	04/13	Question Answering 1		
25	04/18	Question Answering 2		Project Baseline Due

26	04/20	Interpreting NNs
27	04/25	Interactive Learning in NLP
28	04/27	Multilingual NLP
29	05/02	Ethics and Wrap-up
30	05/04	Final Project Presentation

Final Project Due on 05/09

Honor Code

The class enforces the [GMU Honor Code](#) and the [more specific honor code policy](#) special to the Department of Computer Science. You will be expected to adhere to this code and policy.

Note to Students

Take care of yourself! As a student, you may experience a range of challenges that can interfere with learning, such as strained relationships, increased anxiety, substance use, global pandemics, feeling down, difficulty concentrating and/or lack of motivation. All of us benefit from support during times of struggle. There are many helpful resources available on campus and an important part of having a healthy life is learning how to ask for help. Asking for support sooner rather than later is almost always helpful. GMU services are available, and treatment does work. You can learn more about confidential mental health services available on campus at: <https://caps.gmu.edu/>. Support is always available (24/7) from Counseling and Psychological Services: 703-527-4077.

Disabilities

If you have a documented learning disability or other condition which may affect academic performance, make sure this documentation is on file with the [Office of Disability Services](#) and come talk to me about accommodations. I will work with you to ensure that accommodations are provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Services, I encourage you to contact them at ods@gmu.edu.

Last updated on Nov 14, 2021