Overview

CS 410/510: Natural Language Processing
Fall 2021

Instructor: Ameeta Agrawal, ameeta@cs.pdx.edu
TA: Shayan Jalalipour, shayan2@pdx.edu

Lectures: Mondays and Wednesdays, 8:15am – 10:05am
Location: Shattuck Hall ANNEX or Attend Anywhere
Instructor office hours: Mondays right after class (in person), Tue. 3pm (Zoo or by appointment
TA office hours: Fridays 4-5pm, in the fish bowl (seating area just outside th CS department, first floor of the FAB).
Class Zoom: Meeting ID 861 1143 1313 (Monday 09/27

Office Hours Zoom: https://pdx.zoom.us/j/82572777476

Slack Channel:

https://join.slack.com/t/pdx-cs/shared_invite/zt-dfn44ipg-a768KWKxBYDXDchWG3ee9
Students should join the "**nlp-fall2021**" channel once they've joined the workspace.

This course provides an introduction to natural language processing (NLP). V will learn about data-driven models for automatic processing - analyzing, understanding and creating of natural language text. We will start by lookin some traditional text classification algorithms before moving onto some rec machine learning techniques for NLP. We will also learn about some key linguistic phenomena that the NLP algorithms attempt to model. Along the we will look at many interesting NLP applications such as machine translatic summarization and information extraction.

Course Prerequisites. Proficiency in Python. Basic probability, statistics and linear algebra. Familiarity with neural networks is helpful but not required.

linear algebra. Familiarity with neural networks is helpful but not required.

Reference Texts. Most of the assigned readings will be from the following tv textbooks, both of which are available online at the authors' websites:

- [SLP] Dan Jurafsky and James H. Martin. Speech and Language Proces (3rd ed. draft)
- [NLP] Jacob Eisenstein. Natural Language Processing

>> (optional) Natural Language Processing - A machine learning perspective, Yue Zhang and Zhiyang Teng

Schedule (this will be updated as needed during the course)

Week	Date	Topic	Readings	A
1	**Mon Sep 27**	Introduction [slides] [video]	• [SLP] Ch. 1 • [SLP] Ch. 2.2-2.4 • [NLP] Ch. 1	
1	**Wed Sep 29**	Text Normalization		(D
2	**Mon Oct 04**	Language Modeling, N-grams	• [SLP] Ch. 3 • [NLP] Ch. 6 through 6.2	
2	**Wed Oct 06**	Text Classification Application (Sentiment Analysis), Project Proposals/Research Discussion	• [SLP] Ch. 4 • [SLP] Ch. 5 • [NLP] Ch. 2.1 through 2.6	
	Mon			(

3	Mon Oct 11	Logistic Regression	• [SLP] Ch. 5	
3	Wed Oct 13	Vector Semantics and Embeddings	• [SLP] Ch. 6	
4	Mon Oct 18	Neural networks and neural language models	• [SLP] Ch. 7	
4	Wed Oct 20	Neural networks and neural language models	• [SLP] Ch. 7	
5	Mon Oct 25	Project Discussion		
5	Wed Oct 27	Fairness/Bias	• Semantics derived automatically from language corpora contain human biases • Lipstick on a Pig: Debiasing methods cover up systematic gender biases in word embeddings but do not remove them	

6	**Mon Nov 01**	Fairness/Bias		
6	**Wed Nov 03**	NER and POS tagging	• [SLP] Ch. 6 • [NLP] Ch. 8.1	(D
7	**Mon Nov 08**	Information Extraction	• [SLP] Ch. 17 Information Extraction	
7	**Wed Nov 10**	TBD		(
8	**Mon Nov 15**	Machine Translation, Sequence2Sequence	• [SLP] Ch. 11 • Sequence to Sequence Learning with Neural Networks, 2014	

			• BLEU: a Method for Automatic Evaluation of Machine Translation, 2002	
8	**Wed Nov 17**	Machine Translation, Attention, Transformers	• OpenAI GPT-2 • GPT-3 • BERT	
9	**Mon Nov 22**	Dialog	• [SLP] Ch. 24 Chatbots and Dialogue Systems	
9	**Wed Nov 24**	Question Answering, Natural Language Generation	• The Curious Case of Neural Text *De*generation, 2020 • SQuAD: 100,000+ Questions for Machine Comprehension of Text, 2016	
10	**Mon Nov 29**	TBD		

10	Wed Dec 01	FINAL PROJECT PRESENTATIONS (TBD)		Fi Re (D

Coursework and Grading

- **4 Assignments (40%)**: There were will be 4 assignments (each worth 10 points) containing both written questions and programming parts. Students enrolled in CS 510 will have an additional research-oriented component to the assignments. You may form study groups but the assignments are to be completed and submitted individually through D2L. In other words, your work must be your own.
- **1 Final Project (55%)**: The final project (e.g., extending one of the recent works published at ACL, EMNLP or NAACL) will be chosen by students (in consultation with the instructor). All reports should use the ACL style files (Overleaf template or downloadable LaTeX/Microsoft Word templates). Please note the following deliverables,

    1. **Midterm report (15%)**: introduce your research question, its motivations and novelty, a brief overview of the related literature (2 pages).
    2. **Project presentation (20%)**: an in-class/remote (?) presentation about your project (5 minutes). -- TBD, how to do this in an Attend Anywhere format. Otherwise, these points will be adjusted towards midterm and final reports.
    3. **Final report (20%)**: a conference-style paper (4-6 pages, not including references) describing your complete project.

- **Class Participation (5%)**: We appreciate your active participation and you can earn credit by attending and interacting during the lectures in a constructive and informed manner, asking and answering questions on the Slack channel, etc.

Course Policies

Late policy: You are allowed a total of 3 late days to use for assignments over the entire semester (no exceptions for project deliverables) -- please

over the entire semester (no exceptions for project deliverables) – please use them wisely. Each late day gives you an extra 24 hours.

D2L: Please check D2L for all announcements and use Slack for all questions as this can also help other students who may have similar questions. All assignments and course grades will also be posted on D2L.

Honor code: Academic integrity is a vital part of the educational experience at PSU. Please see the PSU Student Code of Conduct for the university's policy on academic dishonesty. A confirmed violation of that Code in this course may result in failure of the course. Copying or paraphrasing someone's work (including code) is not allowed and will result in an automatic grade of 0 for the assignment/project in which copying/paraphrasing was done. If you believe you are going to have trouble completing an assignment, please talk to the instructor or TA well in advance of the due date.

CLASSROOM REQUIREMENTS FOR ALL STUDENTS AND FACULTY DUE TO COVID-19

The University has established rules and policies to make the return to the classroom as safe as possible. It is required for everyone to follow all the Return to Campus rules and policies. To participate in this class, PSU requires all students to comply with the following.

MASKS REQUIRED AT ALL TIMES IN CLASSROOM

- Wear a mask or face covering indoors at all times. Your mask or face covering must be properly worn (fully covering nose and mouth and tight-fitting). Mesh masks, face shields, or face covering that incorporates a valve designed to facilitate easy exhalation are not acceptable. **Because a mask must be worn in the classroom, there should be no eating or drinking in the classroom.** If you have a medical condition or a disability that prevents you from wearing a mask or cloth face covering, you must obtain an accommodation from the Disability Resource Center (DRC) to be exempt from this requirement.

- CDC, State, and County guidance does not limit class size for in-person instruction or require physical distancing.

VACCINATION

- Be vaccinated against COVID-19 and complete the COVID-19 vaccination attestation form. Those students with medical or nonmedical exemptions or who will not be on campus at all must complete the process described on the “COVID-19 Vaccine Exemption Request Form” to establish those exemptions.

HEALTH CHECK, ILLNESS, EXPOSURE, OR POSITIVE TEST FOR COVID-19

- Complete the required self-check for COVID-19 symptoms before coming to campus each day.
- If you are feeling sick or have been exposed to COVID-19, do not come to campus. Call The Center for Student Health and Counseling (SHAC) to discuss your symptoms and situation at 503-725-2800. They will advise you on testing, quarantine, and when you can return to campus.
- If you test positive for COVID, report your result to SHAC and do not come to campus. SHAC will advise you on quarantine, notification of close contacts, and when you can return to campus.
- Please notify me (i.e. your instructor), should you need to miss a class period for any of these reasons so that we can discuss strategies to support your learning during this time.
- If I become ill or need to quarantine during the term, either I or the department chair will notify you via PSU email about my absence and how course instruction will continue.

FAILURE TO COMPLY WITH ANY OF THESE RULES

As the instructor of this course, the University has given me the authority to require your compliance with these policies. If you do not comply with these requirements, I may ask you to leave the classroom, or I may need to cancel the class session entirely.

In addition, failure to comply with these requirements may result in a referral to the Office of the Dean of Student Life to consider charges under PSU's Code of Conduct. A student found to have violated a university rule

(or rules) through the due process of student conduct might face disciplinary and educational sanctions (or consequences). For a complete list of sanctions, see Section 14 of the Student Code of Conduct & Responsibility

sanctions, see Section 14 of the Student Code of Conduct & Responsibility.

GUIDANCE MAY CHANGE

Please note that the University rules, policies, and guidance may change at any time at the direction of the CDC, State, or County requirements. Please review the University's main COVID-19 Response webpage and look for emails from the University on these topics.

Student Services

Disability Access Statement

If you have, or think you may have, a disability that may affect your work in this class and feel you need accommodations, contact the Disability Resource Center to schedule an appointment and initiate a conversation about reasonable accommodations. The DRC is located in 116 Smith Memorial Student Union, 503-725-4150, drc@pdx.edu, https://www.pdx.edu/disability-resource-center/

Safe Campus Statement

Portland State University desires to create a safe campus for our students. As part of that mission, PSU requires all students to take the learning module entitled Creating a Safe Campus: Preventing Gender Discrimination, Sexual Harassment, Sexual Misconduct and Sexual Assault. If you or someone you know has been harassed or assaulted, you can find the appropriate resources on PSU's Enrollment Management & Student Affairs: Sexual Prevention & Response website at http://www.pdx.edu/sexual-assault

Student Food Security

Any student who has difficulty affording groceries or accessing sufficient food to eat every day, or who lacks a safe and stable place to live, and believes this may affect their performance in the course, is urged to contact Committee for Improving Student Food Security for support at foodhelp@pdx.edu. Furthermore, please notify the professor if you are

comfortable in doing so. This will enable her to provide any resources that she may possess.