San Francisco State University
Computer Science Department
Spring 2024

Course Syllabus

Course Number:	CSC 620 & CSC 820
Course Title:	Natural Language Technologies
Number of Credits:	3
Course Level:	Senior undergraduates and Graduates
Class hours:	TTh 12:30 to 1:45PM
Class location:	Hensill Hall 206
Instructor:	Dr. Anagha Kulkarni (ak@sfsu.edu)
Office hours:	Tuesdays 4:30 to 5:30 PM over Zoom (link on Canvas)
Prerequisite:	Consent of the instructor & CSC 413 (grade C or better)
Textbook:	"Speech and Language Processing (3rd Edition)" by Jurafsky & Martin. Free digital copy is available at: https://web.stanford.edu/~jurafsky/slp3/ "Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit" by Steven Bird, Ewan Klein, and Edward Loper https://www.nltk.org/book/
Course Website	Canvas: https://sfsu.instructure.com/
Course Objectives:	**The objectives of this course are to:** 1. introduce the fundamental concepts and techniques of Natural Language Processing (NLP). 2. provide in-depth understanding of the computational properties of natural language and the commonly used algorithms for processing language data. 3. illustrate the pervasiveness of natural language technologies in current times through real-world applications. 4. provide practical experience through hands-on exercises.

Learning Outcomes: At the end of this course students will have:

1. a thorough understanding of NLP fundamentals.
2. a good grasp of the internals of a modern NLP system.
3. a broad understanding of the applications of NLP.
4. the skillset to develop or extend complex NLP systems.

Class Outline

Tuesdays:

12:30 to 12:40PM — Weekly quiz
12:45 to 1:45PM — Lecture

Thursdays:

12:30 to 12:45PM — Review quiz from Tuesday
12:45 to 1:45PM — Lecture

Grading:

Rubric:

For Undergrads (CSC 620):

1. Lecture attendance & active participation: 15%
2. Weekly quizzes: 24% (12 quizzes x 2pts)
3. Homework assignments: 39% (13 assignments x 3pts)
4. Final exam: 22%

1. Lecture attendance & active participation:
   - We won't use the last few rows of chairs in this class, unless we have to.
   - No electronic devices during class.

2. Weekly quizzes
   - Tuesdays' class will start with a short quiz that is based on last week's reading material + lectures' contents.
   - Thursdays' class will start by reviewing students' responses to the quiz questions.
   - We will skip the quiz during these 3 weeks: Jan 29th, Feb 19th, Apr 8th

3. Homework assignments:
   - Typically a coding blog post or Jupyter notebook with functioning codebase that tackles an NLP topic will be provided.
   - You will be expected to read, understand, comment, and expand on what is given.
   - For some assignments you will develop the complete program.

- These are short assignments with work scope of 1-week, and are intended to provide hands-on experience and programming practice for NLP.
- Typically, assigned on Thursday after class, and due by next Thursday before class.

For Graduate students (CSC 820): The grading rubric is same as above but the final exam will be different. For graduate students the exam will consist of more in-depth questions, and questions based on the following seminal paper:
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I., 2017. Attention is all you need. *Advances in neural information processing systems*, *30*.

Recommended Time Allocation: 9 hours / week (at least)

Weekly time allocation for Undergrads:

1. Lectures: 1.5 hours
2. Reading and prepping for weekly quiz: 3.5 hours
3. Doing homework assignments: 4 hours

Weekly time allocation for Graduate Students:

1. Lectures: 1.5 hours
2. Reading and prepping for weekly quiz: 4 hours
3. Doing homework assignments: 3 hours
4. Reading & researching the paper: 0.5 hours

The final letter grade for the course will be calculated using the following distribution:

Highest %	Lowest %	Letter Grade		Highest %	Lowest %	Letter Grade
100	93	A		76.99	73	C
92.99	90	A-		72.99	70	C-
89.99	87	B+		69.99	67	D+
86.99	83	B		66.99	63	D
82.99	80	B-		62.99	60	D-
79.99	77	C+		59.99	0	F

Tentative Schedule of Topics:

Week	Readings	Topics
1	Ch 1 (From Edition 2)	Introduction to NLP and its applications
2	Ch 2	Basics of Text Processing. Regular Expressions. Edit Distance.
3	Ch 3	Language Modeling and N-Grams
4 & 5	Ch 4	Naïve Bayes Classification Evaluation
6, 7, 8	Ch 5	Logistic Regression Cross Validation & Statistical Significance Testing
9	Ch 6	Vector Semantics
10 & 11	Ch 7	Neural Nets and Neural Language Models
12 & 13	Ch 9	Sequence Processing with Recurrent Networks

Course Policies:

Class Attendance: Lecture attendance is mandatory. Student has to have at least 90% lecture attendance to pass the course. For planned and unavoidable absences student should secure permission from the instructor ahead of time by email. In extenuating situations students should send an explanation via email to the instructor as soon as possible after the missed lecture.

Late homework submissions: Late submissions are not accepted unless a prior exception has been secured from the instructor.

Academic Integrity: I take academic integrity very seriously and I have no tolerance for cheating and plagiarism. I have adopted 0-strike policy for my courses which means that the first infraction will lead to failing the course.

Students with disabilities: Students with disabilities who need reasonable accommodations are encouraged to contact the instructor. The Disability Programs and Resource Center (DPRC) is available to facilitate the reasonable accommodations process. The DPRC is located in the Student Service Building and can be reached by telephone (voice/TTY 415-338-2472) or by email (dprc@sfsu.edu).

SF State fosters a campus free of sexual violence including sexual harassment, domestic violence, dating violence, stalking, and/or any form of sex or gender discrimination. If you disclose a personal experience as an SF State student, the course instructor is required to notify the Title IX Coordinator by completing the report form available at http://titleix.sfsu.edu, emailing vpsaem@sfsu.edu or calling 338-2032. To disclose any such violence confidentially, contact:
The SAFE Place - (415) 338-2208; http://www.sfsu.edu/~safe_plc/
Counseling and Psychological Services Center - (415) 338-2208
http://psyservs.sfsu.edu/
For more information on your rights and available resources: http://titleix.sfsu.edu