

Topics in Natural Language Processing

CS 273

Fall 2026 In Person 3 Unit(s) 08/19/2026 to 12/07/2026 Modified 09/02/2026

Contact Information

Instructor(s): Faranak Abri

Office Location: MQ 213

Telephone: 4089244844

Email: faranak.abri@sjsu.edu

Office Hours: Tuesdays 4:30 PM - 5:30 PM

Class Days/Time: Tuesdays – Thursdays 3:00 PM - 4:15 PM

Classroom: MQH 422

Course Description and Requisites

Variable topics in Natural Language Processing. Content may include regression/classification model evaluation, text processing, spell correction, language models, sentiment classification, keyword search, and other NLP-based systems.

Prerequisite(s): Graduate standing. CS 156 or CS 171 or instructor consent.

Grading: Letter graded.

Classroom Protocols

- Students are expected to assist in maintaining a classroom environment that is conducive to learning. Inappropriate behavior in the classroom that leads to the distraction of others shall not be tolerated under any circumstances.
- Instruction will begin at or within several minutes of the official published start time for the course. Please make sure that cell phones, beepers, and texting devices are turned off during the entire scheduled class time. Excessive audible discussions with fellow students are prohibited so that others are not disturbed. If any subject matter is not understood, please do not hesitate to ask for clarification.

If an extended response is necessary to remove doubts, then a request to follow up outside of scheduled classroom instruction time might be made.

- Per [University Policy S12-7 \(https://www.sjsu.edu/senate/docs/S12-7.pdf\)](https://www.sjsu.edu/senate/docs/S12-7.pdf), course material developed by the instructor is the intellectual property of the instructor and cannot be shared publicly without permission. Students may not publicly share or upload instructor-generated material for this course such as exam questions, lecture notes, or homework solutions without the instructor's consent. This includes unauthorized recording or posting of recordings of lectures. Students who record, distribute, or post these materials will be referred to the Student Conduct and Ethical Development office. These policies are designed to protect student privacy and ensure academic integrity.
- If a student is caught cheating on a homework assignment, the student will receive a 0 on that assignment. If a student is caught cheating on an exam, the student will receive an F. The instructor must report any incidents of cheating or plagiarism to the University per [University Policy F15-7 \(https://www.sjsu.edu/senate/docs/F15-7.pdf\)](https://www.sjsu.edu/senate/docs/F15-7.pdf).
- Attendance for the first two sessions is required. Important class information, including policies and class schedule, will be explained. Students who do not attend the first sessions will be removed from the course.
- Exam 1 and Exam2 will be held in person. So, students have to make sure they attend the exam sessions in person.

Program Information

Diversity Statement - At SJSU, it is important to create a safe learning environment where we can explore, learn, and grow together. We strive to build a diverse, equitable, inclusive culture that values, encourages, and supports students from all backgrounds and experiences.

Course Learning Outcomes (CLOs)

Upon successful completion of this course, students will be able to:

- CLO1 Design regression and classification learners (models) and evaluate the goodness of their designed Models based on overfitting/underfitting and bias/variance.
- CLO2 Process textual data, and design/evaluate ML models to classify textual data
- CLO3 Analyze the main applications of NLP and implement different NLP projects.
- CLO4 Write technical reports about their projects and present their findings.

Course Materials

Textbooks (Optional)

- Speech and Language Processing (An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition). Dan Jurafsky and James H. Martin.
 - Free access: <https://web.stanford.edu/~jurafsky/slp3/> (<https://web.stanford.edu/~jurafsky/slp3/>)
- Pattern recognition and machine learning. Christopher M. Bishop; 2006

- Machine Learning. Sergios Theodoridis; 2015
- Foundations of Statistical Natural Language Processing. Christopher Manning and Hinrich Schütze
- Other online resources

☰ Course Requirements and Assignments

- Two exams 40% = First exam 20% + Second exam 20%
 - Note: Exam 1 and Exam2 will be hold in-person. So, students have to make sure they will attend the exam sessions in-person.
- Three projects 60% (Each project includes one or more subprojects, written reports in LaTeX format) = First project 10% + Second Project 20% + Capstone project (required presentation) 30%
 - grading rubrics for Project 1 and Project 2: Implementing all mentioned steps (0.6) and clear report (0.4)
 - **Capstone project:** Students must choose a project for their final project that will require an adequate amount of teamwork and then carry it out. The capstone project is a mid-scale project that can be composed of smaller projects, and students work in groups to accomplish their project.
 - The Capstone project proposal is sent to the instructor for approval before starting to work on it. The proposal includes the description of the NLP project, putting it out in one page's worth of detail.
 - Briefly describe the dataset(s) you plan to work on.
 - List the steps you want to take for implementation.
 - Establish what the project's results will be.
- grading rubrics for Capstone project: Project proposal (0.1), implemented project (0.5), Report (0.2), presentation (0.2)
- Bonus points = Presenting research work for specific topics approved by the instructor.

"Per [University Policy S16-9](http://www.sjsu.edu/senate/docs/S16-9.pdf), (<http://www.sjsu.edu/senate/docs/S16-9.pdf>) success in this course is based on the expectation that students will spend, for each unit of credit, a minimum of 45 hours over the length of the course (normally three hours per unit per week) for instruction, preparation/studying, or course-related activities, including but not limited to internships, labs, and clinical practices. Other course structures will have equivalent workload expectations as described in the syllabus."

Final Examination or Evaluation

This is a project-based course, and the final examination is done by the capstone project and the written report in LaTeX format and the presentation (30%).

✓ Grading Information

A+ = 97% or higher	A = 93%–96%	A- = 90%–92%
B+ = 87%–89%	B = 83%–86%	B- = 80%–82%
C+ = 77%–79%	C = 73%–76%	C- = 70%–72%
D+ = 67%–69%	D = 63%–66%	D- = 60%–62%

A+ = 97% or higher	A = 93%–96%	A- = 90%–92%
F = below 60%		

University Policies

Per [University Policy S16-9 \(PDF\)](http://www.sjsu.edu/senate/docs/S16-9.pdf) (<http://www.sjsu.edu/senate/docs/S16-9.pdf>), relevant university policy concerning all courses, such as student responsibilities, academic integrity, accommodations, dropping and adding, consent for recording of class, etc. and available student services (e.g. learning assistance, counseling, and other resources) are listed on the [Syllabus Information](https://www.sjsu.edu/curriculum/courses/syllabus-info.php) (<https://www.sjsu.edu/curriculum/courses/syllabus-info.php>) web page. Make sure to visit this page to review and be aware of these university policies and resources.

Course Schedule

Topics in Generative AI, Fall 2026, Course Schedule

There might be some changes to the course schedule due to the class circumstances.

Week	Dates: Tu, Th	Topics, Readings, Assignments, Deadlines
Week 0	Aug 20	Syllabus
Week 1	Aug 25, Aug 27	Review of evaluation of ML/NLP models, working with Overleaf and LaTeX
Week 2	Sep 1, Sep 3	Review of the evaluation of ML/NLP models P1.1
Week 3	Sep 8, Sep 10	Review of the evaluation of ML/NLP models P1.2
Week 4	Sep 22, Sep 24	Text Preprocessing: regular expressions, stemming, lemmatization, stop words, text Normalization, tokenization P2.1
Week 5	Sep 29, Oct 1	Edit Distance, word and sentence similarity, and cosine similarity P2.2

Week 6	Oct 6, Oct 8	Exam 1, Vector Semantics and Embeddings, Word senses and
Week 7	Oct 13, Oct 15	N-grams language models, evaluation, and smoothing. P2.3
Week 8	Oct 20, Oct 22	Text classification and Naïve Bayes, Sentiment classification Capstone project proposal
Week 9	Oct 27, Oct 29	Chatbots and Dialogue Systems Capstone project proposal
Week 10	Nov 3, Nov 5	RNN, LSTM
Week 11	Nov 10, Nov 12	Transformers, large language models
Week 12	Nov 17, Nov 19	Exam 2 , Capstone project presentations
Week 13	Nov 24, Nov 26	Capstone project presentations
Week 14	Dec 1, Dec 3	Capstone project presentations

No Final Exam

Other important dates:

Sep 15: Last Day to Drop Courses without an Entry on Student's Permanent Record (D)

Fall 2026 calendar:

https://www.sjsu.edu/provost/docs/2026-27_Final%20Draft.pdf