**San José State University**

# Department of Computer Science CS 286-81, Advanced Topics in Computer Science (Topics in NLP), Spring 2023

## Course and Contact Information

Instructor(s): Faranak Abri

Office Location: MQ 213

Telephone: 4089244844

Email: faranak.abri@sjsu.edu

Office Hours: Fridays 2PM-4PM (Zoom or in-person by appointment)

Class Days/Time: TuTh 9:00AM - 10:15AM

Classroom: online (check canvas)

**Prerequisites:**

CS 171 or CS 156 or equivalent

## Course Description

*This is the first course in natural language processing (NLP), and after finishing it, students can proceed on to more advanced materials. In this course, we will review the fundamentals of Machine Learning and also the fundamentals of natural language processing.*

## Course Format

## The course is delivered Online.

### Canvas

It is the student’s responsibility to check canvas regularly. For help with using Canvas see [Canvas Student Resources page](http://www.sjsu.edu/ecampus/teaching-tools/canvas/student_resources).

## Course Learning Outcomes (CLO)

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

* *CLO1 Understand the steps of solving a generic supervised regression or classification problem.*
* *CLO2 Design regression and classification learners(models) and evaluate the goodness of their designed Models.*
* *CLO3 Understand how to work with textual data and design/evaluate ML models*
* *CLO4 Learn about the main applications of NLP*
* *CLO5 Practice their skills for implementing a project, writing reports and presenting their findings*

## Textbook (Optional)

*Pattern recognition and machine learning. Christopher M. Bishop; 2006*

*Machine Learning. Sergios Theodoridis; 2015*

*Speech and Language Processing (An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition).Dan Jurafsky and James H. Martin.*

*Foundations of Statistical Natural Language Processing. Christopher Manning and Hinrich Schütze*

*Online resources*

### Technology

* Students are required to have an electronic device (laptop or tablet). If you do not have access to an electronic device, SJSU has a free equipment loan program available for students ([link](https://sjsuequipment.getconnect2.com/)).

## Course Requirements and Assignments

* *Two exams 30% = First exam 15% + Second exam 15%,*
* *Three Projects 70% (Each project includes one or more subprojects, written reports in LaTex format and upon the request presentations) = First project 15% + Second Project 20% + Capstone project 35%*
* *Bonus points = Presenting research work for specific topics*

“Per [University Policy S16-9,](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 practica. 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 (35%).

## Grading Information

*Extra credit questions may be included in assignments and exams. For late submissions, grade deduction will be considered.*

|  |  |  |
| --- | --- | --- |
| *Grade* | *Points* | *Percentage* |
| *A plus* | *960 to 1000* | *96 to 100%* |
| *A* | *930 to 959* | *93 to 95%* |
| *A minus* | *900 to 929* | *90 to 92%* |
| *B plus* | *860 to 899* | *86 to 89 %* |
| *B* | *830 to 859* | *83 to 85%* |
| *B minus* | *800 to 829* | *80 to 82%* |
| *C plus* | *760 to 799* | *76 to 79%* |
| *C* | *730 to 759* | *73 to 75%* |
| *C minus* | *700 to 729* | *70 to 72%* |
| *D plus* | *660 to 699* | *66 to 69%* |
| *D* | *630 to 659* | *63 to 65%* |
| *D minus* | *600 to 629* | *60 to 62%* |

“This course must be passed with a C- or better as a CSU graduation requirement.”

## Classroom Protocol

* *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 (Links to an external site.)*](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 instructor 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 a 0. A second incident of cheating will result in the student receiving an F in the course. The instructor must report any incidents of cheating or plagiarism to the University per*[*University Policy F15-7 (Links to an external site.)*](https://www.sjsu.edu/senate/docs/F15-7.pdf)*.*

## University Policies

* Per [University Policy S16-9](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 [Syllabus Information web page](https://www.sjsu.edu/curriculum/courses/syllabus-info.php) (https://www.sjsu.edu/curriculum/courses/syllabus-info.php). Make sure to visit this page to review and be aware of these university policies and resources.
* University policy F69-24 at <http://www.sjsu.edu/senate/docs/F69-24.pdf> states that “Students should attend all meetings of their classes, not only because they are responsible for material discussed therein, but because active participation is frequently essential to ensure maximum benefit for all members of the class. Attendance per se shall not be used as a criterion for grading.” However, attendance will be required in order to complete and submit many in-class exercises, quizzes, and exams.
* It is the aim of the faculty of SJSU to foster a spirit of complete honesty and a high standard of integrity. The University Academic Integrity Policy S07-2 at http://www.sjsu.edu/senate/docs/S07-2.pdf requires you to be honest in all your academic coursework. Faculty members are required to report all infractions to the office of Student Conduct and Ethical Development. The attempt of students to present as their own any work that they have not honestly performed will be considered a violation. During quizzes and exams, communication with other individuals via any means is strictly prohibited without the express permission of the instructor. Violations will be met with the full impact of SJSU’s academic integrity policy and procedures.

## Additional Information

* **Religious Holy Day:** A student who intends to observe a religious holyday should make that intention known in writing to the instructor prior to the absence.

* **Students with Disabilities:** Any student who, because of a disability, may require special arrangements in order to meet the course requirements should contact the instructor as soon as possible to make any necessary arrangements. Students should present appropriate verification from Student Disability Services during the instructor’s office hours. Please note instructors are not allowed to provide classroom accommodations to a student until proper confirmation from Student Disability Services has been provided. Presidential Directive 97-03 at https://sjsu.edu/president/docs/PD\_1997-03.pdf requires that students with disabilities requesting accommodations register with the Accessible Education Center (AEC) at <http://www.sjsu.edu/aec>to establish a record of their disability.
* **Basic Needs**: It is hard to succeed in any course if your basic needs are not met. SJSU has resources to assist with food or housing insecurity, emergency financial loans, and other assistance at SJSU Cares. A list of SJSU resources for student success (tutoring and other educational support), basic needs, mental health resources, COVID-19 relief resources, affinity groups on campus, and other helpful resources can be found here ([link )](https://docs.google.com/document/d/1eEPM-aAvYZFnX85OKOTI-s73TZlwolOSDFS_9CFof0U/edit?usp=sharing). [This](https://docs.google.com/document/d/1cM8akFKZ2VyRciczZ374VSNnzXfyr0tLV8bQp5ah5_Q/edit) list is adapted from a list created by Dr. Wilkinson in the biology department. These resources are available in order to help you succeed and make sure your basic needs are met. Please do not hesitate to make use of any available resources. If you have any questions or additions to this list, please reach out to the instructor.

# CS 286-81 / Advanced Topics in Computer Science, Spring 2023, Course Schedule (exam dates will be announced)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Week/session** | **Date** | **Topics, Readings, Assignments, Deadlines** | **Week/session** | **Date** | **Topics, Readings, Assignments, Deadlines** |
| w1/s1 | 1/26 | Introduction | w1/s2 | 1/31 | preprocessing |
| w2/s3 | 2/2 | regression, models, evaluation | w2/s4 | 2/7 | regression, models, evaluation |
| w3/s5 | 2/9 | overfitting, underfitting | w3/s6 | 2/14 | overfitting, underfitting |
| w4/s7 | 2/16 | classification, models | w4/s8 | 2/21 | classification, evaluation |
| w5/s9 | 2/23 | Regular Expressions, part-of speech, stop words, | w5/s10 | 2/27 | Regular Expressions, part-of speech, stop words, |
| w6/s11 | 3/2 | Text Normalization, tokenization | w6/s12 | 3/7 | Text Normalization, tokenization |
| w7/s13 | 3/9 | lemmatization, stemming | w7/s14 | 3/14 | Edit Distance |
| w8/s15 | 3/16 | N-gram Language Models | w8/s16 | 3/21 | bag of words, TF-IDF (Term Frequency - Inverse Document Frequency) |
| w9/s17 | 3/23 | word and sentence similarity and cosine similarity | w9/s18 | 3/28 | Spring Recess |
| w10/s19 | 3/30 | Spring Recess | w10/s20 | 4/4 | dependency parsing |
| w11/s21 | 4/6 | Sentiment Classification | w11/s22 | 4/11 | Named entity recognition |
| w12/s23 | 4/13 | Machine Translation | w12/s24 | 4/18 | Text summarization |
| w13/s25 | 4/20 | keyword search (Information Extraction | w13/s26 | 4/25 | Question  Answering |
| w14/s27 | 4/27 | Capstone project presentation | w14/s28 | 5/2 | Capstone project presentation |
| w15/s29 | 5/4 | Capstone project presentation | w15/s30 | 5/9 | Capstone project presentation |
| w16/31 | 5/11 | Capstone project presentation | Final Exam | No final exam | <https://www.sjsu.edu/classes/final-exam-schedule/fall-2022.php> |