

SI425 Natural Language Processing Course Policy, Fall AY18

Coordinator: Assoc. Prof. Nate Chambers, Room 366, x3-6838, nchamber@usna.edu

Course Description:

This course covers topics for the machine understanding of human languages. Can computers automatically understand and communicate through natural languages like English? Both the intelligence community and the private sector are increasingly interested in mining huge amounts of written text for information. The course will cover algorithms to both learn and interpret language. Hands-on laboratory assignments include topics like author identification, language modeling, information retrieval from social media, email filtering, syntactic parsing, and sentiment analysis. Prereq: IC312.

Credits: 3-2-2

Learning Objectives:

At the end of the course students should be able to:

1. Design and implement language models with smoothing for unseen words.
2. Discuss the key components of a probabilistic parser.
3. Understand basic English syntax and its use in NLP applications.
4. Describe the difference between Naive Bayes and log-linear algorithms.
5. Design and implement probabilistic document classifiers.
6. Describe basic learning algorithms that use bootstrapping techniques.

Student Outcomes:

SI425 contributes to the following outcomes that you should be able to do upon graduation:

1. Analysis. An ability to analyze a problem, and to identify and define the computing requirements appropriate to its solution.
2. Implementation. An ability to design, implement, and evaluate a computer-based solution to meet a given set of computing requirements in the context of the discipline.

Textbook(s):

We will use the unpublished 3rd Edition of Speech and Language Processing, by Jurafsky and Martin. This is a free online source at <https://web.stanford.edu/~jurafsky/slp3/>

If you prefer a physical text, the Second Edition is available for purchase:

Speech and Language Processing. Jurafsky and Martin. Second Edition; Prentice Hall, 2008. (ISBN-13: 978-0131873216)

Yes, we will actually read this!

Extra Instruction:

Extra Instruction (EI) is available and encouraged when your own attempts to understand the subject matter are unsuccessful. However, you must come prepared with specific questions or areas to be discussed (i.e. have read the assigned readings). If you have missed class, get the notes from a classmate. EI is normally available during weekdays by appointment; see the course web page (URL above) for hours of non-

availability. Students may also show up at the instructor's office without appointment, however no expectation of instructor availability should be assumed. Asking questions over e-mail is strongly encouraged, though the reply might request in-person EI as the most effective solution.

Absences

Students are responsible for obtaining any material missed due to an absence. Additionally, students must ensure that their work is submitted by the deadline regardless of other commitments, i.e. duty, sick call, movement orders. Should bona fide emergencies arise, it is the responsibility of the student to coordinate with the instructor IN ADVANCE (emergency leave, hospitalization, SIR, etc.).

Collaboration: The guidance in the Honor Concept of the Brigade of Midshipmen and the Computer Science Department Honor Policy must be followed at all times. See:

www.usna.edu/CS/resources/honor.php

Specific instructions for this course:

- **Labs and Final Project:** You may collaborate on laboratory assignments and the project to the following extent: collaborative **conversations** with regard to syntax, strategies and methods for accomplishing the goal of the labs are encouraged, however design and implementation must be the work of the individual student handing in the final product. Sharing or copying of code is never permitted. Further, **looking at a classmate's code** for the same lab/project is not permitted. In addition, you **must identify** all of those with whom you collaborate in a text file when you submit your assignments. Consult your instructor if you need further clarification.
- **Homeworks:** Collaboration is allowed on homework. The final code or written submission must be entirely your own, but you may discuss approaches and algorithms with your classmates. You may never look at another student's written answers. This restriction also prohibits "checking your answers" by comparing answers side by side.
- **Exams/Quizzes:** Exams will be open notes, but closed book, **unless otherwise specified. Quizzes are closed book/notes.** All work on exams/quizzes must be your own in accordance with USNAINST 1610.3F, USNAINST 1531.53, and COMPCIDEPTINST1531C. These references can be found at <http://www.usna.edu/cs/resources/honor.php>.
- **Online Resources:** You can use online resources to assist your learning. These often contain code examples. You **must identify** in your final code every instance of code you adapted from an outside source. Using external code without citing it is an honor offense.

All collaboration and outside sources should always be cited. The same rules apply for giving and receiving assistance. If you are unsure whether a certain kind of assistance or collaboration is permitted, you should assume it is not, work individually, and seek clarification from your instructor.

Section Leader

The section leader will call the class to order and record absences and tardiness. If the instructor is absent, do not dismiss class. Instead, contact the department office and lead the class in productive review.

Preparation for Lab

Each student must have their personal lecture notes with them for every class session that meets in the lab. Students without their own notes are unprepared and must retrieve them before they can begin work.

Classroom Conduct:

Students are expected to uphold all professional standards while in class. Proper uniforms shall be worn, and proper language shall be used. Sleeping in class is prohibited. Stand in the back of the room if you are falling asleep.

Late Policy:

Penalties for late submission of graded work may vary among courses or from semester to semester, but they will be the same for all sections of a given course. For *this* course:

Late Assignments and Floating Grace Days

Unless otherwise specified, assignments are due **one minute before lab or class on the due date** (electronically). Assignments with paper copies must be turned in at the start of lab/class on the due date. If the paper copy is later than this, then the whole assignment will be treated as if submitted when the paper copy arrives.

Weekend days count as full late days. An assignment due on Friday is 3 days late if turned in on Monday.

You are encouraged to turn everything in on time like the responsible adult that you are. However, unexpected events do happen, so you have **5 floating grace days** to use during the semester. You may spread these out over any number of labs (up to 4 grace days for any single lab). For instance, you may use 3 grace days on one lab, and 2 grace days on another. After using all of your grace days, you will receive a 0 (zero) on any late assignment thereafter. Weekend days count as full late days.

Please note: grace days are intended to flexibly handle things like illnesses, injuries, and stressful circumstances. You shouldn't have to worry when these things happen. This is your safety net. However, if you use up your 5 days for "trivial" reasons, and then you fall ill, please consider what you're asking before pleading for extra late time.

Electronic Submission Time	Penalty Days
By the due date and time	None
One minute after due time	1 day
23 hours and 59 minutes after due time	1 day
24 hours and 1 second after due time	2 days
etc.	n days

Grading:

	6-Week Grade	12-Week Grade	Final Grade
6-Week Exam	40%	25%	10%
12-Week Exam		25%	10%
Quizzes/Homework/Participation	20%	10%	5%
Labs	40%	40%	35%
Final Project			20%
Final Exam			20%

For the Final Project, pairs of students will independently investigate an approved topic, and build a significant system that learns and solves an open NLP problem. Grades will be based on depth of content, coding challenge, and experimental results. Both students might not receive the same grade.

Exams

The 6 and 12-week exams will primarily focus on the recent material. The final exam will be comprehensive. If a make-up exam is required, inform the instructor at least 1 week in advance. Expect the exams to challenge your understanding of the underlying principles involved. Exam questions will not be simple rewrites of earlier homework assignments, but will instead be entirely different questions that require in-depth understanding to solve.