CS 585: Natural Language Processing

Fall 2023 Syllabus

Course Description

This course is about how to build algorithms and systems that analyze unstructured natural language texts and extract useful information and knowledge from them. Students should expect to gain familiarity with the most common types of natural language processing (NLP) tasks, including text classification, sequence labeling, and structure prediction—and to learn appropriate frameworks for performing these tasks. The course will cover the technical methodology in sufficient detail to allow students to apply these frameworks in an informed way and to make current research accessible.

Prerequisite

CS 430 Introduction to Algorithms with a min. grade of C

Date and Location

Siegel Hall 118 [map]
Tuesday & Thursday 8:35 am - 9:50 am

Instructor

Sonjia Waxmonsky, Ph.D.
Email: swaxmonsky@iit.edu
Office hours (Virtual): Friday 9:30 - 10:30 am
Office hours Zoom link:
https://iit-edu.zoom.us/j/83567906737?pwd=a1ZGRk9tZWdzRVlmZWJHWHJRaXdudz09
Other office hour times: by appointment only via email

Teaching Assistants

TBA when finalized by Computer Science department
TA Office hours: TBA

Grading Policies

Homework Assignments: 5 assignments (10% each → 50% total)
Midterm Exam (20%)
Final Exam (30%)

Late Homework Policy

Homework assignments will be due by 11:59 pm Central Time on the posted due date. Students are encouraged to start homework early and make multiple iterative submissions on Blackboard; only the last homework submission will be graded.

Assignments not received by that time may be submitted up to 24 hours later, with a **50% penalty**. Homework will **not be accepted more than 24 hours after the due date**.

Exceptions to this policy for medical reasons will only be granted with the permission of the Associate Chair of the Computer Science department. Please direct requests for medical exceptions to this late policy, and supporting documentation, to the Associate Chair.

Course Discussion

Students are encouraged to share questions and comments in the Blackboard discussion groups. Please direct general-interest questions on homework and course materials to online course discussion so that all students may benefit from the information.

Textbook and Reading Materials

• M&S: Foundations of Statistical Natural Language Processing, by Christopher Manning and Hinrich Schutze, MIT Press, 1999.
https://github.com/won1k/cs287/blob/master/Manning_Schuetze_StatisticalNLP.pdf
• E-NLP: Introduction to Natural Language Processing, by Jacob Eisenstein, MIT Press, 2019.
https://github.com/jacobeisenstein/gt-nlp-class/tree/master/notes
• UNIXWiki: https://en.wikibooks.org/wiki/Guide_to_Unix/Commands/Text_Processing

Academic Integrity and Student Conduct

Violations of the University Academic Integrity policy will not be ignored. Penalties include reduced or no credit for submitted work, a failing grade in the class, a note on your official transcript that shows you were punished for cheating, suspension, expulsion and revocation of already awarded degrees. The university requires that should I implement any of these penalties, I must report the matter to the Dean's office.
Code of Academic Honesty: https://web.iit.edu/student-affairs/handbook/fine-print/code-academic-honesty