CSC 485/2501 Introduction to Computational Linguistics
Course Information

Instructor: Jingcheng (Frank) Niu
Prof. Gerald Penn oversees the essays for graduate students.

Lectures:

Course/Section	Time	Room
CSC485H1F LEC0101	MWF 10-11	MWF: ES B142.
CSC485H1F LEC0201	MWF 12-13	M: MP 137, WF: ES B142.
CSC2501H1F LEC0101	MWF 12-13	M: MP 137, WF: ES B142.

Office Hours: MWF 11-12, or by appointment.
Email: csc485-2024-09@cs.toronto.edu
Teaching Assistants: Jinman Zhao (Lead), Zixin Zhao (A1), Yushi Guan (A2), Jinyue Feng (A3), Bindu Dash and Devan Srinivasan.

Course Website, Piazza, Quercus and Other Important Links

- **Course Website**: `https://www.cs.toronto.edu/~niu/csc485/`.

- **Piazza**: Please monitor the Piazza discussion board. Your teaching assistants will be monitoring it and all assignment clarifications will be posted there. For non-confidential inquiries, consult the Piazza forum first.

- **Quercus**: Non-public materials will be posted on Quercus. You should be automatically enrolled if you're registered for the course.

- **MarkUS**: Please submit your code to the course MarkUS page.

Textbooks

- Jurafsky and Martin, Speech and Language Processing.

- Bird, Klein and Loper, Natural Language Processing with Python.

Both books are available online for free, so the physical copy is optional.

Evaluation   There will be three assignments, one quiz per week, a bonus course content survey and no final exam. For CSC 2501 students, there will also be 5 essays to write based upon assigned readings.

Component	CSC485	CSC2501
Assignment 1	30%	25%
Assignment 2	30%	25%
Assignment 3	30%	25%
Quizzes	10%	10%
Course Content Survey	Bonus 1%	Bonus 1%
Essay 1∼5		3% Each

- **Quizzes**: There will be a weekly quiz (except during reading week), which will be announced after one of the lectures. You will need to complete the quiz on Quercus. You will earn 1 mark for correctly completing the quiz, 0.5 marks for completing it on time but incorrectly, and 0 marks if you miss the quiz. Your final quiz grade will be based on your 10 best scores.

- **Course Content Survey**: We will be conducting a course content survey to gather your feedback and learn which sub-topics of the course you are most interested in. The course content will be adjusted based on the

survey results. By completing the survey before the end of Reading Week (Sunday, November 3, 2024, by 11:59 p.m.), you will earn an 1% bonus mark. Survey link: `https://forms.gle/AeT3QFGnT8CLdZF77`.

- **Lateness**: No late assignments will be accepted except in case of documented medical or other emergencies.

Prerequisites For undergraduates, STA237/247/255/257 and CSC209, but CSC311 and CSC324/384 are strongly recommended. Engineering students may substitute APS105/APS106/ESC180/CSC180 for the CSC209 requirement, although experience with the Unix operating system is strongly recommended, and/or ECE302/STA286/CHE223 /CME263/MIE231/MIE236/MSE238/ECE286 for the statistics requirement. Note that the University's automatic registration system checks for prerequisites: even if you have registered for the class, you will be dropped from it unless you had satisfied the prerequisite before you registered or you had received a prerequisite waiver. For advice, contact the Undergraduate Office on the fourth floor of the Bahen Centre or the instructor.

Remark Requests Requests for remarking an assignment must be made within seven (7) days of the return of the marked assignment via the remark request online form (link TBA).

Requests for remarking will be reviewed by the head TA or instructor if deemed necessary, consult with the grading TA, who originally mark the assignment. If the grading TA determines that the original grade was too high, the student's grade may be lowered. If the grading TA determines that the original grade was too low, the student's grade will be adjusted accordingly. Once all the remarks are completed, they will be released back to the students. The decision on the remark will be final and no further requests for a remark will be considered. Any or all other departmental and university policies on remarking applies as well.

Policy on Collaboration, Using Online Resources, AI Writing Assistance and Plagiarism Collaboration on and discussion of quiz content is encouraged. No collaboration on homeworks or essays is permitted. The work you submit must be your own. No student is permitted to discuss or share homeworks with any other student from either this or previous years.

Posting solutions, materials, or handouts from assignments, quizzes, or essays to any public forum (including but not limited to Reddit, GitHub, and GitLab) is strictly prohibited. The use of any unauthorized online materials is also forbidden. Submitting any code or writing that is not your own constitutes an academic offence.

The use of AI writing assistance (ChatGPT, Copilot, etc) is allowed only for refining the English grammar and/or spelling of text that you have already written. Submitting any Python, Prolog, TRALE and any other code generated or modified by any AI assistants is strictly prohibited.

Failure to observe this policy is an academic offence, carrying a penalty ranging from a zero on the homework to suspension from the university. See Academic integrity at the University of Toronto.

Important Dates

Wed, Sep 4	First lecture.
Mon, Sep 16	Essay 1 due (CSC 2501).
Mon, Sep 30	Essay 2 due (CSC 2501).
Thu, Oct 3	Assignment 1 due (CSC 485/2501).
Wed, Oct 16	Essay 3 due (CSC 2501).
Mon, Oct 14	Thanksgiving holiday (no lectures or tutorial).
Oct 28 - Nov 1	Reading Week (no lectures or tutorial).
Sun, Nov 3	Bonus credit deadline for the Course Content Survey (CSC 485/2501).
Mon, Nov 4	Last day to drop course (CSC 485/2501).
Thu, Nov 7	Assignment 2 due (CSC 485/2501).
Mon, Nov 11	Essay 4 due (CSC 2501).
Mon, Nov 25	Essay 5 due (CSC 2501).
Mon, Dec 2	Last lecture.
Wed, Dec 4	Assignment 3 due (CSC 2501/485).

Tentative Syllabus   There will be minor changes on the schedule based on progress and class feedback.

- Week 1. Introduction
- Week 2. Dependency Grammar & Parsing
- Week 3. The Turing Test
- Week 4. Syntax and Interpretation
- Week 5. Lexical semantics, vector semantics and Word Sense Disambiguation
- Week 6. Large Language Models
- Week 7. Chart Parsing
- Week 8. Parsing with Features
- Week 9. Statistical parsing
- Week 10. Supertagging
- Week 11. Unsupervised Parsing
- Week 12. Question Answering
- Week 13. Anaphora Resolution