Course Syllabus

Course Information

Course Number/Section   CS6320.003
Course Title   Natural Language Processing
Term   Fall 2023
Days & Times   Mon & Wed 10:00 am-11:15 am
Location   HH 2.502 (or Teams)

Professor Contact Information

Professor   Xinya Du
Email Address   xinya.du@utdallas.edu
Office Location & Hours   ECSS 3.227, Friday 1:30-2:30pm
Class Web Page   Link

Teaching Assistant Information

Teaching Assistant   Ruosen Li, Liqiang Jing
Email Address   TBD
Office Hours   Monday 2-3pm on Teams.

Course Prerequisites, Co-requisites, and/or Other Restrictions
CS 5343 — Algorithm Analysis and Data Structures

Course Description
This graduate course provides an introduction to the field of natural language processing. We will cover core NLP tasks, such as language modeling and parsing, as well as complex applications, such as question answering, machine translation, and information extraction. We will discuss a variety of machine learning approaches to NLP, including statistical and neural approaches.

Learning Objectives/Outcomes
- Ability to compute n-gram statistics and to perform smoothing
- Ability to perform part-of-speech tagging
- Knowledge about semantic analysis, including lexical semantics and word sense disambiguation
- Knowledge of machine learning & deep learning-based models and their applications to NLP applications (such as text categorization, information extraction, information retrieval, summarization, question answering)

Textbooks and Materials

- (Recommended) Dan Jurafsky & James H. Martin, *Speech and Language Processing (3rd Edition)* — https://web.stanford.edu/~jurafsky/slp3/ **(J&M)**.
- Yoav Goldberg, Neural Network Methods for Natural Language Processing, 2017.

Assignments & Academic Calendar (*Tentative, Due Date: all 11:59pm CST*)

Day	Date	Topic	Assignments	Recommended Readings
~~Mon~~	~~Aug 21~~	~~Logistics & Introduction~~		
Wed	Aug 23	-		
Mon	Aug 28	Language Modeling		J&M 3
Wed	Aug 30	-		
Mon	Sep 4	No class, Labor Day		
Wed	Sep 6	Text Classification		J&M 4, 7 (.2)
Mon	Sep 11	-	a1	
Wed	Sep 13			J&M 8 (.1--.4)
Mon	Sep 18	Sequence Labeling		
Wed	Sep 20	-		
Mon	Sep 25	-		
Wed	Sep 27	Neural Networks		J&M 7 (.1--.4), Primer
Mon	Oct 2	Computation Graphs, BackProp		J&M 7, Intro to Computation Graphs
Wed	Oct 4	Vector Semantics & Word Embeddings		J&M 6, word2vec explained
Mon	Oct 9	-	a1 due, a2 out	
Wed	Oct 11	Recurrent Neural Networks		J&M 9 (.1--.6), Karphacy15
Mon	Oct 16	-		
Wed	Oct 18	Recurrent Neural Networks, Seq2Seq, Attention		J&M 9, J&M 10 (.2, .3), Luong 15
Mon	Oct 23	Midterm Review		
Wed	Oct 25	**Midterm (in ECSS 2.102)**		
Mon	Oct 30	-		
Wed	Nov 1	Self-Attention, Transformers, BERT		**illustrated Transformer**, Annotated Transformer, Paper
Mon	Nov 6	NP coreference resolution		J&M 11, The Illustrated BERT, ELMo, and co
Wed	Nov 8	Question Answering	a2 due	
Mon	Nov 13	Guest Speaker (Kalpesh Krishna from Google Bard Team)		J&M 17
Wed	Nov 15	QA/NLG		J&M 21
Mon	Nov 20	No class, Fall Break.		

Day	Date	Topic	Assignments	Recommended Readings
~~Mon~~	~~Aug 21~~	~~Logistics & Introduction~~		
Wed	Nov 22	No class, Fall Break.		
Mon	Nov 27	QA/NLG/Prompting		
Wed	Nov 29	Prompting and (or RLHF)		
Mon	Dec 4	Final Review		
Wed	Dec 6	Final Exam (MC 2.410)		
~~Final Exam Week (Dec 9 -- 15)~~		~~TBD~~		

The descriptions and timelines contained in this syllabus are subject to change at the discretion of the instructor.

Grading

- Assignments/Projects 45%
- Midterm 20%
- Final 30%
- Participation & quiz 5%

Grading Scale
- A   93 or above
- A-   90-93
- B+   87-90
- B   83-87
- B-   80-83
- C+   77-80
- C   70-77
- F   70 or below

Course Policies

Late Work Policy	If the homework is turned in after the deadline, the grade for the homework shall be reduced by 20% for the first 24 hours, 50% for the next 24 hours, and will not be accepted after 48 hours.
Make-up exams	There will be no make-up exams, projects, or homework.
Class Attendance	Regular class participation is expected. Students who fail to participate in class regularly are inviting scholastic difficulty.
Comet Creed	This creed was voted on by the UT Dallas student body in 2014. It is a standard that Comets choose to live by and encourage others to do the same: *"As a Comet, I pledge honesty, integrity, and service in all that I do."*
UT Dallas Syllabus Policies and Procedures	The information contained in the following link constitutes the University's policies and procedures segment of the course syllabus. Please review the catalog sections regarding the credit/no credit or pass/fail grading option and withdrawal from class. Please go to https://go.utdallas.edu/syllabus-policies for these policies.

Class Materials
The professor may provide class materials that will be made available to all students registered for this class as they are intended to supplement the classroom experience. These materials may be downloaded during the course, *however, these materials (including exams) are for registered students' use only and should not be posted publicly. Classroom materials may not be reproduced or shared with those not in class, or uploaded to other online environment*s except to implement an approved Office of Student AccessAbility accommodation. Failure to comply with these University requirements is a violation of the Student Code of Conduct.

Academic Integrity

The faculty expects from its students a high level of responsibility and academic honesty. Because the value of an academic degree depends upon the absolute integrity of the work done by the student for that degree, it is imperative that a student demonstrate a high standard of individual honor in his or her scholastic work.

Scholastic dishonesty includes, but is not limited to, statements, acts or omissions related to applications for enrollment or the award of a degree, and/or the submission as one's own work or material that is not one's own. As a general rule, scholastic dishonesty involves one of the following acts: cheating, plagiarism, collusion and/or falsifying academic records. Students suspected of academic dishonesty are subject to disciplinary proceedings.

Plagiarism, especially from the web, from portions of papers for other classes, and from any other source is unacceptable and will be dealt with under the university's policy on plagiarism (see general catalog for details). This course will use the resources of turnitin.com, which searches the web for possible plagiarism and is over 90% effective.

Withdrawal from Class

The administration of this institution has set deadlines for withdrawal of any college-level courses. These dates and times are published in that semester's course catalog. Administration procedures must be followed. It is the student's responsibility to handle withdrawal requirements from any class. In other words, I cannot drop or withdraw any student. You must do the proper paperwork to ensure that you will not receive a final grade of "F" in a course if you choose not to attend the class once you are enrolled.

These descriptions in this syllabus are subject to change at the discretion of the Professor, and Welcome to the course!