Course Syllabus

Course Information

Course Number/Section	CS 4395.002
Course Title	Human Language Technologies
Term	Spring 2024
Days & Times	Mon & Wed 11:30am - 12:45pm
Location	ECSW 1.315

Professor Contact Information

Professor	Xinya Du
Email Address	xinya.du@utdallas.edu
Office Location & Hours	Online, Friday 3-4pm

Teaching Assistant Information

Teaching Assistant	Ruosen Li, Mian Zhang
Email Address	mian.zhang@utdallas.edu
Office Hours	via Teams

Course Prerequisites, Co-requisites, and/or Other Restrictions

CS3345: Data Structures and Introduction to Algorithmic Analysis
CS3341: Probability and Statistics in Computer Science
Familiarity with basic probability, algorithms, (differential) multivariable calculus, and linear algebra is also assumed.

Course Description

This 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.

Textbooks and Materials

- (Recommended) Dan Jurafsky & James H. Martin, *Speech and Language Processing (*3^{rd} *Edition)* — https://web.stanford.edu/~jurafsky/slp3/ **(J&M)**.
- Machine Learning by Tom Mitchell

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

Day	Date	Topic	Assignments	Recommended Readings
~~Mon~~	~~Jan 15~~	Logistics & Introduction		
Wed	Jan 22	-		
Mon	Jan 27	Language Modeling		J&M 3
Wed	Jan 29	-	a1 out	
Mon	Feb 3	Text Classification		J&M 4, 7 (.2)
Wed	Feb 5	-		J&M 8 (.1--.4)
Mon	Feb 10	-		
Wed	Feb 12	Sequence Labeling		
Mon	Feb 17	-		
Wed	Feb 19	Neural Networks		J&M 7 (.1--.4), Primer
Mon	Feb 24	~~Computation Graphs, BackProp~~		~~J&M 7, Intro to Computation Graphs~~
Wed	Feb 26	Vector Semantics & Word Embeddings	a1 due	
Mon	Mar 3	**Pytorch / Python tutorial**		J&M 6, word2vec explained
Wed	Mar 5	-		J&M 9 (.1--.6), Karphacy15
Mon	Mar 10	Recurrent Neural Networks		
Wed	Mar 12	Recurrent Neural Networks, Seq2Seq, Attention	a2 out	J&M 9, J&M 10 (.2, .3), Luong 15
Mon	Mar 17	Spring break		
Wed	Mar 19			
Mon	Mar 24	-		J&M 9 (.1--.6), Karphacy15
Wed	Mar 26	Self-Attention, Transformers, BERT		**illustrated Transformer**, Annotated Transformer, Paper
Mon	Mar 31	-		J&M 11, The Illustrated BERT, ELMo, and co
Wed	Apr 2	**Midterm Review**		
Mon	Apr 7	**Midterm**		
Wed	Apr 9	Final Project Feasibility Check-In (on Teams)		
Mon	Apr14	LLM / prompting	a2 due	
Wed	Apr 16	Natural Language Generation (NLG)		
Mon	Apr 21	Invited Talk (Dr. Majumder)		

Wed	Apr 23	Question Answering (QA)	quiz
Mon	Apr 28	Final presentations (groups 1--6)	
Wed	Apr 30	Final presentations (group 7 -- 12)	
Mon	May 5	Final presentations (group 13--18)	
Wed	May 7	Final presentations (group 19-24)	
	May 12		final project report due (11:59pm)

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

Homework

We expect you to try solving each problem set on your own. However, if you get stuck on a problem, I encourage you to collaborate with other students in the class, subject to the following rules:

- You may discuss a problem with any student in this class, and work together on solving it. This can involve brainstorming and verbally discussing the problem, going together through possible solutions, but should not involve one student telling another a complete solution.
- Once you solve the homework, you must write up your solutions on your own, without looking at other people's write-ups or giving your write-ups to others.
- In your solution for each problem, you must write down the names of any person with whom you discussed it. This will not affect your grade.

Grading

- Assignments/Projects 38%
- In-class quiz and course feedback 2%
- Midterm (closed-book) 30%
- Final project presentation 10% -- example 1, example 2
- Final project report 20% -- examples reports

Grading Scale

- A+ >= 96
- A 93-96
- A- 90-93
- B+ 87-90
- B 83-87
- B- 80-83
- C+ 77-80
- C 74-77
- C- 65-75
- F 65 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. Later homeworks will **NOT** be accepted except in extreme circumstances or those permitted by university policy (e.g., a religious holiday). All such exceptions MUST be cleared in advance of the due date if possible.
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!