Natural Language Course Syllabus

Course Information

Course Prefix and Number: CS7650
Credit Hours: 3
Instructor: Goyal, Kartik
Semester, Year: Fall 2026

Course Description

Natural Language Processing and Understanding is a multi-disciplinary field that focuses on computational methods to analyze and manipulate natural human language with relevance to artificial intelligence, machine learning, computer science, and linguistics. This course focuses on the fundamental principles of Natural Language Processing and Language Modeling. Large language model products and agents like ChatGPT, Claude, and Gemini are ubiquitously used today and this course will focus on underlying algorithms, methods , and techniques involved in building such language models and other computational methods to broadly analyze language. This course will cover machine learning principles, neural architectures for processing language, training and inference algorithms, techniques for adaptation and manipulation of language models, and other relevant research topics in linguistics, large-scale analysis of unstructured natural language data, and language modeling. The students will develop skills related to the topics above by engaging in an open-ended project, programming homeworks, theory assignments, and classroom activities.

Course Objectives and Learning Outcomes

By enrolling in this course, students will:

1. Understand underlying technical principles behind development of language models and natural language processing tools.
2. Gain confidence in building computational tools for language and language models in addition to using them.
3. Learn task-specification and evaluation methodologies for NLP tools including language models.
4. Familiarize themselves with current research questions and state-of-the-art in NLP/language modeling.

Required Course Materials

Textbooks: Due to the nature of the course, instead of relying completely on a standard textbook, a lot of the material in this course will involve reading research papers, tutorials, and lecture notes. The following textbooks will be used to supplement the course material:

a) SLP3.: Daniel Jurafsky and James H. Martin. Speech and Language Processing. Drafts of some chapters of the 3rd edition are freely available online..
b) INLP.: Jacob Eisenstein. Introduction to Natural Language Processing. 2019. A 2018 draft is available online.

Compute resources: The programming assignments will ask you to implement state-of-the-art natural language processing algorithms using neural networks. For this purpose we will use Pytorch, and you will require access to GPUs. We encourage you to explore Google Colab., which provides easy access to GPUs. If using Google Colab, we **highly** recommend signing up for Colab Pro (once you start working on the part of the homework that uses Pytorch). This costs $10 / month, which is roughly equivalent to the cost of a textbook over the course of the semester. This will provide a better experience working on the homework assignments by giving you access to better GPUs, etc. Additionally, you will have access to PACE-ICE and you should be able to work on your homeworks and project using PACE-ICE.
You are welcome to use other GPU resources to complete the assignments if you choose, but we cannot provide support for this. *You will need to submit your code and program output in Jupyter Notebook format that runs on Colab Pro.*

Canvas: This class will use Canvas to deliver course materials. Gradescope will be used for homeworks, assignments, and project. Piazza will be used for communication and class discussion.

Grading Policy and Weighting

Grading has following components:

a) Semester Project including proposal, literature survey, progress report, project presentation, and final report -- 20%,
b) 4 Programming Homeworks -- 40%,
c) 3 Non-programming Problem Sets -- 30%,
d) Preparatory concept-check homework (hw0) and problem set (ps0) -- 5%
e) Class Participation (see attendance policy below) -- 5%

Your grade will be assigned as a letter grade, with the following score threshold guarantees (i.e., more than 90% will definitely be an A)

a) A -- greater than 90%
b) B -- greater than 80%
c) C -- greater than 70%
d) D -- greater than 60%
e) F -- 60% and below

Attendance Policy

This course meets twice weekly and all students are expected to attend the sessions. Three special sessions pertaining to two project presentations and a project clinic require attendance and will involve recording of attendance. The attendance for lectures will not be recorded explicitly, however lecture participation contributes towards the final grade as demonstrated by student engagement on Piazza and completion of random pop-quizzes.

Additional Criteria for Successful Completion of the Course

Students will benefit from prerequisite knowledge Probability and Statistics, Linear Algebra, Multivariate Calculus, Programming/Python experience, and prior exposure to fundamental machine learning and deep learning concepts.

You will be afforded some late days for deliverables except the quizzes and the semester project deliverables, as detailed in the policy described on the Canvas homepage. Any concessions exceeding the late days will only be made under extenuating circumstances. In case of such emergencies, please contact the Dean of Students. office (see here. for rules). The Dean of Students is equipped to verify emergencies and pass confirmation on to all your classes. For consistency, we ask all students to do this in the event of an emergency. Do not send any personal/medical information to the instructor or TAs; all such information should go through the Dean of Students.

Academic and Research Honesty/Integrity Statement

Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. Review the and the , especially .Student Code of ConductAcademic Honor CodeAppendix A: Graduate Addendum to the Academic Honor Code

Students are expected to perform research in an ethical and responsible manner. All Doctoral and Master's Thesis students are required to take the , and it is expected that students abide by the principles taught in that training while performing research Responsible Conduct of Research training

Any student suspected of cheating or plagiarizing on a quiz, exam, or assignment will be reported to the Office of Student Integrity, who will investigate the incident and identify the appropriate penalty for violations.

Allegations of scientific or scholarly misconduct are handled in accordance with the procedures outlined by the .Policy for Responding to Allegations of Scientific or Other Scholarly Misconduct

Core IMPACTS

Not applicable

Accommodations for Students with Disabilities

If you are a student with learning needs that require special accommodation, as soon as possible to make an appointment to discuss your special needs and to obtain an accommodations letter. Please also e-mail me as soon as possible in order to set up a time to discuss your learning needs. contact the Office of Disability Services

Student-Faculty Expectations: acceptable student conduct

At Georgia Tech, we believe that it is important to strive for an atmosphere of mutual respect, acknowledgement, and responsibility between faculty members and the student body. articulates some basic expectations that you can have of me and that I have of you. Additional information for research-related work is given in . In the end, simple respect for knowledge, hard work, and cordial interactions will help build the environment we seek. Therefore, I encourage you to remain committed to the ideals of Georgia Tech while in this class. The Student-Faculty ExpectationsThe Expectations of Advisors and Advisees