

Computer Understanding of Natural Language (CAP 6640)

Textbook Documents will be provided.

Lecture TuTh 4:30-5:45pm, HEC 119, <https://ucf.zoom.us/j/91739664976>

Instructor Dr. Liqiang Wang Office: HEC 322 (only available on Zoom)
Office Hours: TuW 1:30pm-2:50pm, <https://ucf.zoom.us/j/92473596471>
E-Mail: liqiang.wang@ucf.edu

Prerequisites Students are expected to have basic knowledge of Machine Learning or Artificial Intelligence, and be familiar with Python programming.

Course Goal and Topics

This is an advanced course of natural language processing, mainly focusing on recent advances in the discipline. The course, in essence, covers both principles and advanced topics in learning (i.e., machine learning, deep learning) based approaches to natural language understanding. By the end of this course, every student should be ideally able to answer questions concerning the following aspects of natural language processing, and gain the theoretical foundations and application contexts of:

- Natural language understanding/processing and historical overview.
- Embedding techniques, including word vectors representation and word senses
- Backpropagation in language model training/tuning
- Neural networks and their role in language modeling and generation
- Linguistic structure and applications of dependency parsing
- The probability of a sentence and recurrent neural networks for language models
- Vanishing gradients, their solutions, and fancy RNN structures
- Applications of machine translation, seq2seq learning, and attention models
- Convolutional networks and their applications to natural language processing
- Information from parts of words; subwords and character-level models
- Modeling contexts of use; contextual representation and pretrained models
- Transformers and more on advanced pretrained language models
- Natural language generation and applications
- Reference in language and language models; coreference resolution

- Multitask learning and tree recursive neural networks
- Safety, bias, and fairness in natural language modeling and understanding

Attendance Each student is encouraged to attend lecture in classroom. Due to the COVID-19, we will provide Zoom live and recorded video for these students who have difficulty in attending class in person. If you cannot come to classroom to attend class, please let me know and get my approval.

Grading Grades will be based on the percentage listed below and a curve at the end of the semester. A plus or minus (e.g., A, A-, B+, B) grade will be used.

Assignments	35%
Project	30%
Midterm Exam	15%
Comprehensive Exam	20%

Assignments Homework assignments will be posted on Webcourses@UCF and announced in class. Assignments should be finished individually if without specific team-work requirement. All source code and readme files should be zipped into a single file and submitted on Webcourses@UCF. Late submission will have a penalty of 20% of the score each day.

Paper reading assignments

Some Assignments will be reading assignments. In each of those assignments, each student will be expected to read, summarize, and critique a research paper on the topics being discussed in the class. Each summary is expected to include the following aspects:

1. A summary statement: this statement should be of no less than 400 words and no more than 500 words. The summary statement should highlight the problem statement, the main technique, and the main findings of the paper. Copying the abstract or parts of the paper, as with the above similarity policy, will result in automatic 0 for this assignment. The summary statement should be nontrivial and highlight at a technical level the contribution of the work.
2. Main strengths: describe concisely 3 strengths of the technical contribution in technical terms. Each of those strengths should be no less than 70 words and no more than 100 words (total of up to 300 words).
3. Main weaknesses: describe concisely 3 strengths of the technical contribution of the paper in technical terms. Each of those weaknesses should be no less than 70 words and no more than 100 words (total of up to 300 words).

Project A large component of the grade is dedicated for a project. Project will be conducted by 1-2 students. The project will be on a topic in the NLP's application area, with a preference to AL, ML, and security related topics/integrations. The following is the grade distribution of the project.

- Project proposal 2% (Jan 31, 2021)
- Project milestone 5% (Feb 26, 2021)
- Project presentation 10% (Apr 14, 19, and 21, 2021)
- Project report 18% (April 26, 2021)

Submissions guideline for project elements:

- **Proposal:** the proposal has to be up to 300 words and must include the following information: 1) title of the project, 2) what is the problem the project tries to address, 3) what technique you expect to use for addressing the problem, what outcomes you expect to get as a result. The topic of the proposal has to be discussed and preapproved beforehand with the professor (via webcourse direct messages). Try to find a partner for the project during the first day of the classes if you have determined you would take the class. The submission of the proposal will be due only 1 week after the start of the classes (per the schedule above).
- **Project milestone:** the project milestone has to be exactly two (2) page on the ACM template of 2 columns, highlighting a concise and final problem statement, technique, initial results, and initial discussion.
- **The presentation** has to be up to 12 slides covering the same items in the final report (below).
- **Final report:** the final report has to be exactly six (6) pages + as many pages as needed for references on the ACM template of 2 columns. The final report must include the following (in order): a concise abstract, a clear introduction, a concise problem statement, related work, technique, evaluation, discussion, and conclusion.

Exams In this course, you will have one midterm exam and one final. No make-up exams will be given. No incomplete will be given in this class. Please take that into account as the important guideline when considering whether this course is right for you.

Revision I may revise any aspect of this syllabus at any time.

Core UCF Policy Statements

The Center for Academic Integrity (CAI) defines academic integrity as a commitment, even in the face of adversity, to five fundamental values: honesty, trust, fairness, respect, and responsibility. From these values flow principles of behavior that enable academic communities to translate ideals into action. <http://academicintegrity.org/> .

UCF Creed: Integrity, scholarship, community, creativity, and excellence are the core values that guide our conduct, performance, and decisions.

- **Integrity:** I will practice and defend academic and personal honesty.
- **Scholarship:** I will cherish and honor learning as a fundamental purpose of my membership in the UCF community.
- **Community:** I will promote an open and supportive campus environment by respecting the rights and contributions of every individual.
- **Creativity:** I will use my talents to enrich the human experience.
- **Excellence:** I will strive toward the highest standards of performance in any endeavor I undertake.

The following definitions of plagiarism and misuse of sources come from the Council of Writing Program Administrators and have been adopted by UCF's Department of Writing & Rhetoric.

Plagiarism

In an instructional setting, plagiarism occurs when a writer deliberately uses someone else's language, ideas, or other original (not common-knowledge) material without acknowledging its source. This definition applies to texts published in print or on-line, to manuscripts, and to the work of other student writers.

Misuse of Sources

A student who attempts (even if clumsily) to identify and credit his or her source, but who misuses a specific citation format or incorrectly uses quotation marks or other forms of identifying material taken from other sources, has not plagiarized. Instead, such a student should be considered to have failed to cite and document sources appropriately.

Responses to Academic Dishonesty, Plagiarism, or Cheating

UCF faculty members have a responsibility for your education and the value of a UCF degree, and so seek to prevent unethical behavior and when necessary respond to infringements of academic integrity. Penalties can include a failing grade in an assignment or in the course, suspension or expulsion from the university, and/or a "Z Designation" on a student's official transcript indicating academic dishonesty, where the final grade for this course will be preceded by the letter Z. For more information about the Z Designation, see <http://goldenrule.sdes.ucf.edu/zgrade>. For more information about UCF's Rules of Conduct, see <http://www.osc.sdes.ucf.edu/>.

Unauthorized Use of Class Materials

There are many fraudulent websites claiming to offer study aids to students but are actually cheat sites. They encourage students to upload course materials, such as test questions, individual assignments, and

examples of graded material. Such materials are the intellectual property of instructors, the university, or publishers and may not be distributed without prior authorization. Students who engage in such activity are in violation of academic conduct standards and may face penalties.

Unauthorized Use of Class Notes

Third parties may be selling class notes from this class without my authorization. Please be aware that such class materials may contain errors, which could affect your performance or grade. Use these materials at your own risk.

Course Accessibility Statement

The University of Central Florida is committed to providing access and inclusion for all persons with disabilities. This syllabus is available in alternate formats upon request. Students with disabilities who need specific access in this course, such as accommodations, should contact me as soon as possible to discuss various access options. Students should also connect with Student Accessibility Services (Ferrell Commons, 7F, Room 185, sas@ucf.edu, phone (407) 823-2371).

Campus Safety Statement

- Emergencies on campus are rare, but if one should arise in our class, we will all need to work together. Everyone should be aware of the surroundings and familiar with some basic safety and security concepts. In case of an emergency, dial 911 for assistance.
- Every UCF classroom contains an emergency procedure guide posted on a wall near the door. Please make a note of the guide's physical location and consider reviewing the online version at http://emergency.ucf.edu/emergency_guide.html.
- Familiarize yourself with evacuation routes from each of your classrooms and have a plan for finding safety in case of an emergency. (Insert class-specific details if appropriate)
- To stay informed about emergency situations, sign up to receive UCF text alerts by going to my.ucf.edu and logging in. Click on "Student Self Service" located on the left side of the screen in the tool bar, scroll down to the blue "Personal Information" heading on your Student Center screen, click on "UCF Alert", fill out the information, including your e-mail address, cell phone number, and cell phone provider, click "Apply" to save the changes, and then click "OK."
- If you have a special need related to emergency situations, please speak with me during office hours.

Deployed Active Duty Military Students

If you are a deployed active duty military student and feel that you may need a special accommodation due to that unique status, please contact me to discuss your circumstances.