University of Oklahoma
School of Computer Science/Graduate Program
C S 5293-001: Natural Language Processing
Fall 2025 (3 credits)

Instructor: Jie Cao jie.cao@ou.edu
TA: Joshua Wiseman Joshua.T.Wiseman-1@ou.edu

Meetings:
Class: In-person, MWF 12 - 12:50 pm, Carson Engr Ctr 0438
Office Hours:

- Instructor: 12:30-1:30PM, Thursday
- TA: 1:00-2:00PM, Monday

You can reach all the course staff at once via a private post on Discord.

Course Prerequisite: C S 2413 or C S 2414 or C S 5005; C S 2813 or MATH 2513; and a statistics course from the departmentally approved list.

Course Description:
This course covers the fundamental algorithms and methods in natural language processing (NLP). Topics include n-gram models, text classification, statistical models over linguistic structures (e.g., sequences, trees, and graphs), vector semantics, neural networks, large language models, etc. Applications include parsing, information extraction, machine translation, topic modeling, dialog systems, and more. Students will build computational models for different areas of NLP.

Textbook

- Speech and Language Processing, Dan Jurafsky and James H. Martin, 3rd Edition, 2024. Free Book, Available Online at https://web.stanford.edu/~jurafsky/slp3/ (Required for reading material)
- Extra Lecture Notes on Special Topics (Will be shared on Canvas), such as LLMs, Agentic AI.

Learning Outcomes:
You will be able to …

- To understand the problems, algorithms, methods of natural language processing, and their relation to both linguistics and statistics.
    - Understand the basic linguistic phenomena and fundamental NLP tasks
    - Use linguistic feature engineering and manipulate probability formula for statistic NLP models
    - Understand the neural representation, deep learning for NLP, and large language models.
- Be able to design, implement, and evaluate practical NLP systems
    - Analyze text via N-gram language models, and linguistic features.

    - Train and evaluate text classification models and improve them with neural networks.
    - Design and implement basic structured prediction methods for sequence, tree, and graph.
    - Improve model performance with neural modes and pretrained language models.
    - Prompt-engineering and post-training large language models for specific tasks.
- Be able to understand research papers in NLP, contribute initial findings via collaboration.
    - Author a report to survey a research area for your groups' research goal.
    - Design initial experiments to compare existing methods and explore new ideas.

Course Expectations & Grading: Students will be required to complete several in-class **quizzes** throughout the course. There will be also 5 **assignments** that each student should individually complete, as well as a **group project** (at most 3 students per project). The course letter grade will be assigned based on the overall percentage: >= 90 (A), >=80 and < 90 (B), >=70 and < 80 (C), >=60 and < 70 (D), and <60 (F). The allocation of percentages is given below:

	Percentages
Assignments	50%
Course Project	30%
In-Class Paper Quizzes	15%
Participation	5%

Participation includes your contributions to class discussions, whether live (during lectures, TA sessions, or in-class quizzes) or online (through public posts or replies on our discord discussion, including anonymous or private questions). We have found that active participation not only helps you engage more deeply with the material but also has a significant impact on your final grade.

Course Policies

Generative AI Policy

I hope that this course aids you in a creative exploration of Generative AI tools and how they can be used to assist you in accomplishing the goals of this course. In this course, you:

- Should never use any AI tools for **in-class quizzes**, violation will lead to 0 quiz score.
- May use any **free** Generative AI tools for **assignments and project**, such as Copilot, or any Open LLMs. This will allow everyone in the course to have the **same access** to Generative AI. You **should not fully delegate** the code writing to AI. It is better to just use the autocompletion features of AI.
- You **MUST** cite your usage of Generative AI, including the version/date of AI models, any direct quotes or paraphrasing of ideas/content generated by AI, per our class's citation guidelines.
- **Must** provide a statement at the end of the assignment about your Generative AI usage. This statement should be created without the use of Generative AI and should be your

own reflection. This statement must include: the platform(s) and prompt(s) you used, a summary of how Generative AI helped you achieve the learning objectives of the course associated with this assignment, and what additional work you did to verify the output, and to make the work your own.

To implement this policy, we will have ongoing discussions in class about effective use of Generative AI and how you are using it. You'll be asked to reflect frequently about AI usage both in-class and as part of your assignment submission (see assignment instructions for specifics), where you may also be asked to share your prompts, screenshots of your chats with Generative AI tools, and how you used Generative AI to prepare you to, by the end of the semester, independently (without the use of Generative AI) achieve the learning objectives of this course.

- If you have any questions about this policy, please talk with me.
- Consequences for Violating the Generative AI Usage Policy: Deviating from the acknowledgement, citation, and reflection guidelines described here may be considered a violation of the academic integrity policy of this course. Per our usage policy, you will be responsible for accuracy, including appropriately citing and summarizing any articles you find through AI research tools, and thus must read the material you are citing. Submitting data or research that is not real (a risk when overly relying on Generative AI) may result in an academic integrity violation for falsifying information. Additionally, there may be times, such as in-class quizzes, specially constrained assignments, where Generative AI usage is prohibited. Any use of AI in those cases will be considered a violation of the academic integrity policy.
- My Use of Generative AI: I will model appropriate Generative AI usage by clearly disclosing when I use it and why. Expected use cases include: revising quiz or exam questions and responses (with my oversight), drafting case studies or educational games to help connect our course topics to the real-world, using my notes and previous PowerPoints to improve the structure of my lectures so that they are clearer to you all as students, graphic design, revising assignment instructions and rubrics to improve clarity for students, receiving feedback on how I communicate with students, and using AI research tools to find current articles to update our course readings. I will never use Generative AI to grade your work.

Attendance: You are expected to attend or view all class lectures. Participation grades, however, are based on your engagement in lectures, TA sessions, check-in quizzes, and online forums. If you are ill, please notify us in advance so we can arrange an appropriate alternative.

Readings: The course schedule lists a set of readings for each lecture day, mainly from the 3rd Edition of Speech and Language Processing by Dan Jurafsky and James H. Martin. You are responsible for this material before class begins. You don't need to buy the book because the PDF chapters are legally free, and you can download the latest version from the textbook website. Besides the book, for new content about LLM and Agentic AI, we will provide free lecture notes or Jupiter Notebook for reading materials and tutorials.

Assignments: We will have both written and programming assignments. These will be made available through and submitted through Canvas. Programming assignments will be done in Python, along with associated packages and libraries.

Lateness Policy: Assignments are typically due at 12:00pm(noon), an hour before the start of class. It's important to get assignment done on time so that you can follow the subsequent lectures. Thanks. We will use the timestamp on Canvas as the submission time. **Assignments will be accepted up to 24 hours after the deadline** but will be assessed a 10% penalty. That is, if your assignment is late and scores 90, then your actual grade will be 81 = 90 - 9. **Assignments will *not* be accepted 24 hours after the deadline.** We will be strict about this policy: A submit time of 12:01 PM, you will face a 10% penalty! This may sound harsh, but we must draw a line somewhere. **Exceptions**: All submissions are subject to the late day policy stated here. We understand, however, that certain factors may occasionally interfere with your ability to hand in work on time. If that factor is an extenuating circumstance such as a medical condition, we ask you to provide documentation directly issued by the University, and we will try to work out an agreeable solution with you.

Shared Task or Project Proposal: The final project will be done in groups (individual projects require prior approval). It could be selected from a list of given shared task, or you could propose your own research project. **No double dipping projects across multiple classes**. In any case, the data and evaluation methods should be **clearly defined**, **accessible and ready to use.** Three evaluation milestones will be used to grade the project, initial proposal writing, mid-term demo, final report. Please closely sync with Instructor and TA for the progress.

Quizzes: There are **both scheduled quizzes** (around 5-15 minutes for 2-4 questions, graded towards the 15%) **and occasional short unannounced check-ins** (2-3 minutes, 1-2 questions, only count as participation score).

- Scheduled quizzes will be administered in class on paper to assess your understanding of the material. They are **closed book**, but you may use your **paper notes** for reference, no other materials/devices are allowed. The quiz scope and schedule will be announced **at least one week in advance**.
- Unannounced check-ins are primarily for **practice and engagement**. They will **count for participation** credit but will not significantly affect your grade. Please make sure you have a device to make a quick answer within 2-3 minutes.

Exams: No exam for this Fall 2025 class.

Canvas: We'll be using Canvas for our learning platform. Readings, assignments, and general course information will be posted there. Assignment submissions and exams will be done through Canvas as well.

Discord and Private Emails: We'll use Discord for better class discussion and communication. The quickest way to get answers to simple questions is to post them on the corresponding channel. If your question is of a private nature, Discord allows you to send direct messages to the instructors and TA. Matters of personal interest should be directed to email instead of to the forum, e.g., informing the instructors of an extended personal illness. The signup link for our class forum is on the Canvas page.

Incompletes: The grade of "I" is intended for the rare circumstance when a student who has been successful in a class has an unexpected event shortly before the class's end. We will not consider giving a student a grade of "I" unless the following three conditions have been met:

1. It is within two weeks of the end of the semester.
2. The student has a grade of C or better in the class.
3. The reason that the student cannot complete the class is properly documented and compelling.

Classroom Conduct: Because cell phones and laptops can distract substantially from the classroom experience, students are asked not to use either during class except when required as part of a classroom exercise. Disruptions of the class will also not be permitted. In the case of disruptive behavior, we may ask that you leave the classroom and charge you with violating the Student Code of Responsibilities and Conduct. Examples of disruptive behavior include:

- Allowing a cell phone or pager to repeatedly beep audibly.
- Playing music or computer games during class in such a way that they are visible or audible to other class members.
- Exhibiting erratic or irrational behavior.
- Behavior that distracts the class from the subject matter or discussion.
- Making physical or verbal threats to a faculty member, teaching assistant, or class member.
- Refusal to comply with faculty or teaching assistant direction.

Proper Academic Conduct: Feel free to discuss all assignments with the instructors or the TAs. Code (projects and assignment): you may discuss code solutions with other students. However:

1. You may not look at or share code with others. Similar code will zero
2. If you discuss a solution with anyone, you must document their names in your assignment.
3. You must **document this in your code if you use external resources** (e.g., Discord, StackOverflow.com, ChatGPT, or any other AI tools).

Make sure that your computer account is properly protected. Use an appropriate password, and do not give your friends access to your account or computer system. Do not leave printouts, computers, or thumb drives around a laboratory where others might access them. Programming projects will be checked by software designed to detect collaboration. This software is extremely effective and has withstood repeated reviews by the campus judicial processes.
Upon the first documented occurrence of inappropriate collaborative work or of taking a solution from a network resource, the instructors will report the academic misconduct to the Campus Judicial Coordinator. The procedure is documented in the University of Oklahoma Academic Misconduct Code (http://integrity.ou.edu). The provider and receiver of a solution will be treated equally in the misconduct process.

Course Reflection Survey: You'll receive a Course Reflection Survey at the end of each semester for each course that you are enrolled in. I strongly encourage you to complete this survey. Your feedback can help me adjust my class for future semesters to help other students be successful. Your feedback is confidential, and I will only receive it after final grades are due. Course Reflection Survey results may also factor into teaching evaluations and annual performance reviews and are shared with department and program chairs.

Tentative Course Schedule (Fall 2025)
https://www.ou.edu/registrar/academic-records/academic-calendars/fall-2025

Week (M, W, F)	Topics	Assignments (Due Date)
1 (Aug 25-29)	Course Introduction/Tokens, Words, Vocabularies, Lexicons	
2 (Sept 1-5)	Monday: Labor Day Holiday	
	Language Models with N-Grams	
3 (Sept 8-12)	Basic Machine Learning, and Text Classification, Feature Engineering, Logistic Regression	Assignment 1: Tokenization and Ngam Language Model
4 (Sept 15-19)	Symbolic Representation, Language Structures, Structured Prediction (Sequence Labeling, Parsing) Sept 18(Thursday), Job Fair will not impact class	Project Milestone1: Proposal
5 (Sept 22-26)	Neural Representation Learning, Vector Semantics, Embeddings	Assignment 2: Basic Machine Learning and Text Classification
6 (Sept 29-3)	Neural Networks, Recurrent Neural Networks and LSTM	
7 (Oct 6-10)	Machine Translation, Seq2Seq, Attention, Transformer	
8 (Oct 13-17)	Mask Language Model, Self-Supervised Learning, Transfer Learning	Assignment 3: Neural Text Classification and Structured Prediction
9 (Oct 20-24)	LLM, In-context Learning/Prompting	Project Milestone2: Progress-Report, and Mid-term Demo, 10-15 mins demo via two office hours in this week.
10 (Oct 27-31)	Language Generation, Information Retrieval and Retrieval-Augmented Generation	
11 (Nov 3-7)	LLM Pretraining, LLM Post-Training-1: Instruction Finetuning	Assignment 4: Improve previous tasks with neural models (RNN. LSTM, Transformer), pretrained-models (Bert, T5, and More)
12 (Nov 10-14)	LLM Post-Training-2: Reinforcement Learning, Test-time Compute, Chain-of-Thought Reasoning	
13 (Nov 17-21)	Multimodal LLM, Agentic AI and Its Applications	Assignment 5: LLM Post training and Agentic AI
14 (Nov 24-28)	**Monday Class:** Agentic AI Continued	
	Thanksgiving Vacation (Nov 26-30)	
15 (Dec 1-5)	Final Presentation Week	Project Mileston3: Monday/Wednesday/Friday, 10 mins Presentation + 2-3 mins Live QA
16 (Dec 8-12)	Beyond Transformer-based LLMs: State Space Models and Diffusion LLMs	
	No Final Exam	

University Policies

Mental Health Support Services

Support is available for any student experiencing mental health issues that are impacting their academic success. Students can either been seen at the University Counseling Center (UCC) located on the second floor of Goddard Health Center or receive 24/7/365 crisis support from a licensed mental health provider through TELUS Health. To schedule an appointment or receive more information about mental health resources at OU please call the UCC at 405-325-2911 or visit University Counseling Center. The UCC is located at 620 Elm Ave., Room 201, Norman, OK 73019.

Title IX Resources and Reporting Requirement

The University of Oklahoma faculty are committed to creating a safe learning environment for all members of our community, free from gender and sex-based discrimination, including sexual harassment, domestic and dating violence, sexual assault, and stalking, in accordance with Title IX. There are resources available to those impacted, including: speaking with someone confidentially about your options, medical attention, counseling, reporting, academic support, and safety plans. If you have (or someone you know has) experienced any form of sex or gender-based discrimination or violence and wish to speak with someone confidentially, please contact OU Advocates (available 24/7 at 405-615-0013) or University Counseling Center (M-F 8 a.m. to 5 p.m. at 405-325-2911).

Because the University of Oklahoma is committed to the safety of you and other students, and because of our Title IX obligations, I, as well as other faculty, Graduate Assistants, and Teaching Assistants, are mandatory reporters. This means that we are obligated to report gender-based violence that has been disclosed to us to the Institutional Equity Office. This means that we are obligated to report gender-based violence that has been disclosed to us to the Institutional Equity Office. This includes disclosures that occur in: class discussion, writing assignments, discussion boards, emails and during Student/Office Hours. You may also choose to report directly to the Institutional Equity Office. After a report is filed, the Title IX Coordinator will reach out to provide resources, support, and information and the reported information will remain private. For more information regarding the University's Title IX Grievance procedures, reporting, or support measures, please visit Institutional Equity Office at 405-325-3546.

Reasonable Accommodation Policy

The University of Oklahoma (OU) is committed to the goal of achieving equal educational opportunity and full educational participation for students with disabilities. If you have already established reasonable accommodations with the Accessibility and Disability Resource Center (ADRC), please submit your semester accommodation request through the ADRC as soon as possible and contact me privately, so that we have adequate time to arrange your approved academic accommodations.

If you have not yet established services through ADRC, but have a documented disability and require accommodations, please complete ADRC's pre-registration form to begin the

registration process. ADRC facilitates the interactive process that establishes reasonable accommodations for students at OU. For more information on ADRC registration procedures, please review their Register with the ADRC web page. You may also contact them at (405)325-3852 or adrc@ou.edu, or visit www.ou.edu/adrc for more information.

Note: disabilities may include, but are not limited to, mental health, chronic health, physical, vision, hearing, learning and attention disabilities, pregnancy-related. ADRC can also support students experiencing temporary medical conditions.

Religious Observance

It is the policy of the University to excuse the absences of students that result from religious observances and to reschedule examinations and additional required classwork that may fall on religious holidays, without penalty. [See Faculty Handbook 3.15.2]

Adjustments for Pregnancy and Related Issues

Should you need modifications or adjustments to your course requirements because of pregnancy or a pregnancy-related condition, please request modifications via the Institutional Equity Office website or call the Institutional Equity Office at 405/325-3546 as soon as possible. Also, see the Institutional Equity Office FAQ on Pregnant and Parenting Students' Rights for answers to commonly asked questions.

Final Exam Preparation Period

Pre-finals week will be defined as the seven calendar days before the first day of finals. Faculty may cover new course material throughout this week. For specific provisions of the policy please refer to OU's Final Exam Preparation Period policy.

Emergency Protocol

During an emergency, there are official university procedures that will maximize your safety.

Severe Weather: If you receive an OU Alert to seek refuge or hear a tornado siren that signals severe weather.

1. Look for severe weather refuge location maps located inside most OU buildings near the entrances.
2. Seek refuge inside a building. Do not leave one building to seek shelter in another building that you deem safer. If outside, get into the nearest building.
3. Go to the building's severe weather refuge location. If you do not know where that is, go to the lowest level possible and seek refuge in an innermost room. Avoid outside doors and windows.
4. Get in, Get Down, Cover Up
5. Wait for official notice to resume normal activities.

Additional Weather Safety Information is available through the Department of Campus Safety.

The University of Oklahoma Active Threat Guidance

The University of Oklahoma embraces a Run, Hide, Fight strategy for active threats on campus. This strategy is well known, widely accepted, and proven to save lives. To receive emergency campus alerts, be sure to update your contact information and preferences in the account settings section at one.ou.edu.

RUN: Running away from the threat is usually the best option. If it is safe to run, run as far away from the threat as possible. Call 911 when you are in a safe location and let them know from which OU campus you're calling from and location of active threat.

HIDE: If running is not practical, the next best option is to hide. Lock and barricade all doors; turn of all lights; turn down your phone's volume; search for improvised weapons; hide behind solid objects and walls; and hide yourself completely and stay quiet. Remain in place until law enforcement arrives. Be patient and remain hidden.

FIGHT: If you are unable to run or hide, the last best option is to fight. Have one or more improvised weapons with you and be prepared to attack. Attack them when they are least expecting it and hit them where it hurts most: the face (specifically eyes, nose, and ears), the throat, the diaphragm (solar plexus), and the groin.

Please save OUPD's contact information in your phone.

NORMAN campus: *For non-emergencies call (405) 325-1717. For emergencies call (405) 325-1911 or dial 911.*

TULSA campus: *For non-emergencies call (918) 660-3900. For emergencies call (918) 660-3333 or dial 911.*

Fire Alarm/General Emergency

If you receive an OU Alert that there is danger inside or near the building, or the fire alarm inside the building activates:

1. *LEAVE* the building. Do not use the elevators.
2. *KNOW* at least two building exits
3. *ASSIST* those that may need help
4. *PROCEED* to the emergency assembly area
5. *ONCE safely outside, NOTIFY first responders of anyone that may still be inside building due to mobility issues.*
6. *WAIT* for official notice before attempting to re-enter the building.

OU Fire Safety on Campus