

Natural Language Processing

College of William & Mary · Fall 2026

Course Details

Course: Topics in Computer Science: Natural Language Processing
(CSCI 680, Section 04, CRN 13337, 3 Credit Hours)

Dates: 08/26/2026 - 12/15/2026

Time: MWF 10:00AM - 10:50AM

Classroom: Integrated Science Center (ISC) 3280

Website: stephanieschoch.com/nlp

Instructor Information

Name: Stephanie Schoch

Office: ISC 2395

Email: [sschoch\(at\)wm\(dot\)edu](mailto:sschoch@wm.edu)

Office Hours: M 1:00 - 2:00PM

Course Description

This course provides a comprehensive introduction to natural language processing, spanning foundational techniques through large language models. The first half of the course will focus on linguistic and statistical fundamentals before advancing to neural architectures. The second half of the course will focus on the modern LLM pipeline (e.g. pre-training, post-training, and prompting). Students will also engage with advanced topics and emerging areas in NLP (e.g. interpretability, harms and risks of language modeling, data-centric NLP).

Prerequisites

Students should be proficient in Python. Experience with packages such as SciPy, Scikit-learn, and PyTorch is helpful. Students should also have experience with Calculus, Linear Algebra, and Probability & Statistics.

Learning Objectives

By the end of the course, students should be able to:

- Explain foundational linguistic, statistical, and neural concepts underlying modern NLP.
- Implement NLP models and pipelines for various tasks (e.g. text classification, question answering, natural language inference) using modern software frameworks.
- Evaluate NLP systems through experimental design, quantitative metrics, and qualitative error analysis.
- Read and critique current NLP research literature and communicate technical ideas effectively.
- Design and execute an NLP research project involving problem formulation, experimentation, and analysis.
- Assess the limitations, risks, and societal implications of NLP technologies (e.g. LLMs), including consideration for the underlying data.

Required Materials

This course will use a combination of textbooks and supplementary readings. All materials are publicly available as PDFs and will be linked on the course webpage.

Textbooks

- Jurafsky and Martin. “Speech and Language Processing.” 3rd Ed.
- Eisenstein. “Natural Language Processing.”

Grading & Policies

Grading

The course grade will be determined by the following components:

Grading Component	Percentage
Project	50%
Homework	$(3 \times 10\%) = 30\%$
Quizzes	$(3 \times 5\%) = 15\%$
Paper Discussions	5%

Grading questions and inquiries can be asked to the instructor within one week of the grade release date. Grades will be released within two weeks after submission.

Grade Scale

The following scale will be used to map numerical grades to letter grades at the end of the semester:

Grade Range	Letter Grade
≥ 93	A
90 – 92	A-
87 – 89	B+
83 – 86	B
80 – 82	B-
77 – 79	C+
73 – 76	C
70 – 72	C-
67 – 69	D+
63 – 66	D
60 – 62	D-
≤ 59	F

If a student’s grade falls between two ranges (e.g. 92.4), the final letter grade (e.g. A or A-) will be decided by the instructor, with consideration of factors such as attendance, participation, and semester-long improvement. This decision will be final.

Submission and Late Policy

Homework assignments, and other written assignment components (e.g. written project proposal), will be released and collected on Blackboard. All assignments are due at 11:59 PM on the listed due date. Written assignments should be typeset with L^AT_EX (scans of handwritten submissions are not permitted). Students can use generative AI to help convert handwritten solutions to L^AT_EX, however, the student is responsible for ensuring faithfulness to their handwritten responses.

Late Policy Students will be given 1 penalty-free late day for the semester, which can be used on any of the homework assignments. Penalty-free late days are not transferrable between students. Homework and written project deliverables will be accepted up to 3 days late, with 10% of the maximum grade deducted for each day late. For example, if a student submits a homework assignment 2 days late and has a penalty-free late day to use, they can receive a maximum grade of 90%; a student who has previously used their penalty-free late day can receive a maximum grade of 80%. Penalty-free late days cannot be used on any the project deliverables. For example, if a student submits a project deliverable 2 days late, even if they have a penalty-free late day to use, the maximum grade they could receive would be 80%. Extensions on homework assignments will only be granted under special, extenuating circumstances.

Assignments

- **Project (50%):** Students will complete a semester-long small group research project. The project will include the following graded components: 1) project pitch; 2) project proposal; 3) progress report; 4) final project report; 5) final project presentation. Detailed descriptions of each component can be found on the course webpage.
- **Homework (30%):** There will be 3 homeworks over the course of the semester. Homework assignments will primarily be coding-based and will cover topics in the lectures and reading materials.
- **Quizzes (15%):** Over the course of the semester, there will be 3 written quizzes. Quizzes may include short answer and/or multiple-choice questions. Outside of extenuating circumstances, missed quizzes will receive a grade of “0” and cannot be made up.
- **Paper Discussions (5%):** We will have 3 classes dedicated to paper discussions, covering a total of 6 research papers on advanced topics. Students will be graded on their active participation in the discussions. For full credit, students are expected to contribute meaningfully to at least 2 of the 3 paper discussion classes.

Important Dates

Date	Event
August 26	First day of classes
September 4	Last day to add/drop
September 7	Labor Day (No classes)
October 8 - 11	Fall Break
October 26	Last day to withdraw
November 3	Election Day (No classes)
November 23-24	Remote instruction days
November 25-29	Thanksgiving Break (No classes)
December 4	Last day of classes

Course Schedule

This is a tentative schedule that is subject to change as the course progresses.

Week	Dates	Topics
1	08/26, 08/28	Course Overview, Fundamentals, History of NLP; Text Processing
2	08/31, 09/02, 09/04	N-Gram Language Models; Text Classification
3	09/09, 09/11	Text Classification (Cont.); Word Embeddings
4	09/14, 09/16, 09/18	Word Embeddings (Cont.); Project Pitches
5	09/21, 09/23, 09/25	Feedforward Networks; Backpropagation and RNNs
6	09/28, 09/30, 10/02	Seq2Seq; Attention; Language Generation
7	10/05, 10/07	Tranformers
8	10/12, 10/14, 10/16	Hands-On Day; Transformer LMs
9	10/19, 10/21, 10/23	Pre-Training; Scaling Laws & Optimization; Post-Training
10	10/26, 10/28, 10/30	Fine-Tuning & Efficient Adaptation; Prompting & ICL; LLM Agents
11	11/02, 11/04, 11/06	Evaluating LLMs; Responsible Language Modeling; Paper Discussion
12	11/09, 11/11, 11/13	Project Work Day; Interpretability; Paper Discussion
13	11/16, 11/18, 11/20	Robustness & Fairness; Data-Centric NLP; Paper Discussion
14	11/23	Course Outro
15	11/30, 12/02, 12/04	Final Project Presentations

Academic Integrity

Please familiarize yourself with W&M's Honor Code. Plagiarism and other Honor Code violations will result in a final course grade of "F" and referral to the Honor Council.

Accommodations

William & Mary accommodates students with disabilities in accordance with federal laws and university policy. Any student who feels they may need an accommodation based on the impact of a learning, psychiatric, physical, or chronic health diagnosis should contact Student Accessibility Services staff at 757-221-2512 or at sas@wm.edu to determine if accommodations are warranted and to obtain an official letter of accommodation. For more information, please see www.wm.edu/sas.