

The Hong Kong University of Science and Technology
UG Course Syllabus (Fall 2026-27)

Course Title: Introduction to Natural Language Processing

Course Code: COMP 4221

Credits: 3

Pre-requisite: ELEC 2600 OR IEDA 2520 OR IEDA 2540 OR MATH 2411 OR MATH 2421 OR MATH 2431

Meeting Time: Tuesdays and Thursdays, 3:00-4:15 p.m.

Instructor: Prof. Yi R. (May) Fung

Email: yrfung@ust.hk

Office / Consultation: Room 3539; consultation times will be announced on Canvas

Course Description

This course introduces core tasks, models, and algorithms in natural language processing (NLP). Topics span language modeling and text classification, word and sequence representations, linguistic structure, question answering, dialogue systems, and machine translation. Advanced material covers pretrained language models, prompting and instruction tuning, and responsible evaluation. Learning activities combine lectures and readings with programming assignments, examinations, and a course project.

Intended Learning Outcomes (ILOs)

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

1. Explain the principles and assumptions behind fundamental NLP representations, models, and algorithms.
2. Design and implement NLP approaches for practical problems using appropriate data and computational methods.
3. Evaluate NLP systems using suitable metrics, error analysis, and reproducible experimental practice.
4. Critically compare modern NLP techniques and communicate their capabilities, limitations, and responsible-use implications.

Assessment and Grading

This course uses criterion-referenced assessment; grades will not be assigned using a curve. Assessment criteria and task-specific point allocations will be published with each assessment on Canvas.

Assessments

Assignments and project milestones will be submitted through Canvas. Exact coursework deadlines and the midterm arrangement will be announced on Canvas. The final examination will follow the University's central examination schedule.

Assessment Task	Course Grade	Due Date
Individual homework assignments	15%	Dates to be announced on Canvas *

Assessment Task	Course Grade	Due Date
Attendance, in-class QA participation via Mentimeter	10%	Throughout the term *
Group-based course project	25%	Milestones to be announced on Canvas *
Midterm examination	15%	Date to be announced on Canvas *
Final examination	35%	University examination schedule

** Marks for individual coursework tasks will normally be released within two weeks of the due date.*

Mapping of Course ILOs to Assessment Tasks

The mapping below shows where each learning outcome is directly assessed.

Assessed Task	Mapped ILOs	Explanation
Programming assignments	ILOs 1-3	Explain core methods, implement NLP models, and evaluate their behavior on specified tasks.
Course project	ILOs 2-4	Formulate a practical problem, build and evaluate a solution, and communicate limitations responsibly.
Attendance and participation	ILOs 1, 3, 4	Demonstrate preparation and contribute constructively to analysis of methods, results, and responsible use.
Midterm examination	ILOs 1-3	Explain foundational concepts and apply course algorithms and evaluation methods to structured problems.
Final examination	ILOs 1-4	Integrate course concepts, compare approaches, and reason about applications and limitations.

Grading Rubrics

Detailed task-specific point allocations will accompany each assessment on Canvas. Across coursework, evaluation emphasizes technical correctness, appropriate model and metric selection, reproducibility, and clear analysis. The course project additionally assesses problem formulation, experimental rigor, interpretation of results, and communication. Examinations assess conceptual understanding and the ability to apply course methods to unfamiliar problems.

Final Grade Descriptors

Final grades reflect the extent to which the course ILOs have been achieved.

Grade	Short Description	Course-Specific Descriptor
A	Excellent Performance	Demonstrates comprehensive command of NLP concepts and methods. Produces accurate, reproducible work with strong analysis, sound evaluation, and insightful communication that exceeds core requirements.
B	Good Performance	Demonstrates good understanding of the main concepts and can implement and evaluate appropriate NLP approaches. Work is generally correct and well reasoned, with only minor gaps.
C	Satisfactory Performance	Demonstrates adequate knowledge of core material and can address familiar NLP problems. Work shows basic implementation and evaluation competence but contains limitations in depth, rigor, or clarity.

Grade	Short Description	Course-Specific Descriptor
D	Marginal Pass	Demonstrates threshold understanding of essential concepts and partial ability to apply course methods. Work meets minimum learning outcomes but contains notable conceptual or technical weaknesses.
F	Fail	Does not demonstrate sufficient understanding of core NLP concepts or the ability to implement, evaluate, and explain solutions at the required level.

Course AI Policy

Generative AI tools may be used only when the instructions for a specific assessment permit them. Permitted uses may include brainstorming, language editing, code debugging, or exploratory explanations. Students must verify all outputs, remain responsible for submitted work, and include a brief AI-use statement naming the tool and explaining how it was used. Incorporated AI-generated text or code must be cited or clearly annotated. Generative AI is prohibited in examinations and in any assessment designated 'no AI.' Students must not upload restricted materials, personal data, or confidential information to external AI systems. Undisclosed or prohibited use constitutes academic misconduct.

Communication and Feedback

Course announcements, materials, assessment specifications, and marks will be communicated through Canvas. Coursework marks will normally be released within two weeks. Feedback will identify strengths, errors, and areas for improvement. Questions about feedback or marks should be raised within five working days after release.

Resubmission Policy

Resubmission and reassessment are not normally permitted unless stated in the instructions for a specific task. Requests arising from documented exceptional circumstances will be handled under University policy. A mark-review request made within five working days may result in the entire submission being re-evaluated.

Required Texts and Materials

Dan Jurafsky and James H. Martin, *Speech and Language Processing*. Assigned chapters and any supplementary readings will be specified on Canvas, together with lecture notes, assessment instructions, and software guidance.

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to [Academic Integrity | HKUST – Academic Registry](#) for the University's definition of plagiarism and ways to avoid cheating and plagiarism.

Additional Resources

Canvas is the authoritative source for the weekly reading list, lecture materials, assignment updates, and course announcements. Additional research papers and software documentation will be linked as needed.