

The Hong Kong University of Science and Technology

UG Course Syllabus

[Course Title] Introduction to Natural Language Processing

[Course Code] COMP4221

[No. of Credits] 3

[Any pre-/co-requisites]

- Computer science: object-oriented programming and data structures, design, and analysis of algorithms.
- Mathematics: multivariable calculus, linear algebra, probability, stochastic process, and statistics.
- Students are expected to have probability, linear algebra, algorithm design and machine learning background. It is suggested to take an introductory algorithm and an introductory machine learning course before taking this course.

Name: [Instructor(s) Name] Yi R. (May) Fung

Email: [Your Email Address] yrfung@ust.hk

Course Description

This course provides an introduction to the tasks, models and algorithms of natural language processing (NLP) including the fundamental NLP tasks and models (such as language model, text classification, word vectors, sequence labelling, etc.), and several typical NLP applications (such as question answering, dialogue systems and machine translation, etc.), and the linguistic structure annotation (such as constituency parsing and dependency parsing). The course also introduces some advanced topics such as pre-trained language models and their prompt tuning and instruct tuning methods to give the students more perspectives of the NLP industry.

Assessments:

[List specific assessed tasks, exams, quizzes, their weightage]

Assessment Task	Contribution to Overall Course grade (%)
Assignments	20%
Course Project	20%
Midterm Exam	25%
Final Exam	35%

Required Texts and Materials

- Daniel Jurafsky & James H. Martin. *Speech and Language Processing*. Third Edition draft. Draft of January 12, 2022.

[Optional] Additional Resources

- Lee, Raymond ST. *Natural Language Processing: A Textbook with Python Implementation*. Springer Nature, 2023.
- Manning, Christopher, and Hinrich Schutze. *Foundations of statistical natural language processing*. MIT press, 1999.
- Jurafsky, Dan. *Speech & language processing*. Pearson Education India, 2000.