City University of Hong Kong
Course Syllabus

offered by Department of Computer Science
with effect from Semester A 2024/25

Part I Course Overview

Course Title:	Natural Language Processing
Course Code:	CS6493
Course Duration:	One semester
Credit Units:	3 credits
Level:	P6
Medium of Instruction:	English
Medium of Assessment:	English
Prerequisites: *(Course Code and Title)*	CS5481 Data Engineering or CS5487 Machine Learning: Principles and Practice or CS5489 Machine Learning: Algorithms and Applications or CS5491 Artificial Intelligence or SDSC5001 Statistical Machine Learning I or SDSC6001 Statistical Machine Learning II or SDSC8007 Deep Learning
Precursors: *(Course Code and Title)*	Nil
Equivalent Courses: *(Course Code and Title)*	Nil
Exclusive Courses: *(Course Code and Title)*	Nil

Part II Course Details

1. Abstract

This course introduces algorithms and techniques for natural language processing, from computational linguistics for text processing to information extraction for language understanding. The topics include statistical and neural based language modeling, word representation, pretrained language models such as BERT, and large language models, such as GPT and Llama. Basic and advanced natural language processing tasks, such as machine translation, dialog systems, question answering, text classification/labeling/tagging, and knowledge graph will also be introduced.

2. Course Intended Learning Outcomes (CILOs)

(CILOs state what the student is expected to be able to do at the end of the course according to a given standard of performance.)

No.	CILOs	Weighting (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	*A2*	*A3*
1.	Perform statistical and neural processing for syntactic analysis of text properties.			✓	
2.	Perform information extraction for understanding of text semantics.			✓	
3.	Assess the effectiveness of natural language processing and understanding for real-world problems.		✓		
4.	Apply techniques in natural language processing and understanding for innovative applications.		✓	✓	✓
		100%			

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: *Accomplishments*
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

3. Learning and Teaching Activities (LTAs)

(LTAs designed to facilitate students' achievement of the CILOs.)

LTA	Brief Description	CILO No.				Hours/week (if applicable)
		1	2	3	4	
Lecture	Students will engage with the algorithms and techniques for natural language processing and understanding. Related real-world applications such as machine translation and dialog systems will be introduced.	✓	✓	✓		2 hours/ week
Tutorial	Students will work on a different problem set each week during the tutorial sessions, through which they can discover the main characteristics of different natural language processing techniques and integrate them for real-world problems. They will also be invited to present their solutions, and the class will be encouraged to provide comments.	✓	✓	✓		1 hour/ week
Assignments	Students will implement selected natural language processing and understanding approaches, apply these approaches to real-world problems, and interpret the results. In this way, students can analyse the performance of different approaches.	✓	✓	✓		
Project	Students will create a new system design and implement appropriate natural language processing approaches for innovative applications. The students will apply the principles they have learnt from the course for their design.				✓	
Examination	Examination will include questions to assess the capability of students 1) to identify the important features of natural language processing and understanding approaches; 2) to perform critical evaluation of different algorithms for real-world language problems; 3) to modify or design algorithms for challenging applications.	✓	✓	✓		

4. Assessment Tasks/Activities (ATs)

(ATs are designed to assess how well the students achieve the CILOs.)

Assessment Tasks/Activities	CILO No.				Weighting	Remarks
	1	2	3	4		
Continuous Assessment: 60%						
Assignments	✓	✓	✓		30%	Expect to have two assignments with programming elements for algorithm implementation.
Project				✓	30%	Around 1-3 students in a group to finish the project
Examination^: 40% (duration: 2 hours)	✓	✓	✓		40%	
					100%	

^ For a student to pass the course, at least 30% of the maximum mark for the examination must be obtained.

5. Assessment Rubrics

(Grading of student achievements is based on student performance in assessment tasks/activities with the following rubrics.)

Applicable to students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Marginal (D)	Failure (F)
1. Assignments	The ability to implement and assess the effectiveness of different algorithms and techniques.	High	Significant	Moderate	Basic	Not even reaching marginal level
2. Project	The ability and creativity in designing and implementing appropriate algorithms and techniques for innovative applications.	High	Significant	Moderate	Basic	Not even reaching marginal level
3. Examination	The extent to which the students can understand the algorithms and techniques, apply them with appropriate modification or design new solutions for different applications, and evaluate their performances.	High	Significant	Moderate	Basic	Not even reaching marginal level

Applicable to students admitted from Semester A 2022/23 to Summer Term 2024

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B)	Marginal (B-, C+, C)	Failure (F)
1. Assignments	The ability to implement and assess the effectiveness of different algorithms and techniques.	High	Significant	Moderate to Basic	Not even reaching marginal level
2. Project	The ability and creativity in designing and implementing appropriate algorithms and techniques for innovative applications.	High	Significant	Moderate to Basic	Not even reaching marginal level
3. Examination	The extent to which the students can understand the algorithms and techniques, apply them with appropriate modification or design new solutions for different applications, and evaluate their performances.	High	Significant	Moderate to Basic	Not even reaching marginal level

Part III Other Information (more details can be provided separately in the teaching plan)

1. Keyword Syllabus

(An indication of the key topics of the course.)

NLP basics: language models, word representations (distributional representation and contextualized representation), recurrent neural networks, attention mechanism, transformers, pretrained language models, and large language models.

NLP task examples: machine translation, dialog systems, question answering, text classification/tagging/ labeling, knowledge graph, and other advanced topics.

2. Reading List

2.1 Compulsory Readings

(Compulsory readings can include books, book chapters, or journal/magazine articles. There are also collections of e-books, e-journals available from the CityU Library.)

1.	Daniel Jurafsky and James H. Martin, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 3rd edition (online), 2020.
2.	Christopher D. Manning, Hinrich Schutze, *Foundations of Statistical Natural Language Processing*, MIT Press, 1999.
3.	Jacob Eisenstein, *Natural Language Processing*, online, 2018.

2.2 Additional Readings

(Additional references for students to learn to expand their knowledge about the subject.)