

**City University of Hong Kong
Course Syllabus**

**offered by Department of Computer Science
with effect from Semester B 2018/19**

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)	CS5286 Algorithms and Techniques for Web Searching or CS5487 Machine Learning: Principles and Practice or CS5489 Machine Learning: Algorithms and Applications
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, tagging and parsing, name entities and relation extraction, and semantic representation and labeling. Computational models for language applications such as machine translation and dialog systems 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 self-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. Teaching and Learning Activities (TLAs)

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

Teaching pattern:

Suggested lecture/tutorial/laboratory mix: 2 hrs. lecture; 1 hr. tutorial.

TLA	Brief Description	CILO No.				Hours/week (if applicable)
		1	2	3	4	
Lecture	The lecture will focus on the algorithms and techniques for natural language processing and understanding. Related real-world applications such as machine translation and dialog systems will be introduced.	✓	✓	✓		
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.	✓	✓	✓		
Assignments	The 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	The 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: <u>50%</u>						
Assignments	✓	✓	✓		30%	Expect to have two assignments with programming elements for algorithm implementation.
Project				✓	20%	Can be in groups of 2–3 students per project.
Examination [^] : <u>50%</u> (duration: 2 hours)	✓	✓	✓		50%	
					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.)

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

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.)

Basics: text normalization, vector semantics, regression, recurrent network.

Words: n-gram language models, neural language models, word sense disambiguation, word2vec representation, word sentiment.

Grammar: Part-of-speech tagging, context-free grammars, syntactic parsing, dependency parsing.

Information extraction: Name entities recognition, relation extraction, template filling, co-reference resolution.

Applications: Text categorization, machine translation, task-oriented dialog agents, chat bot.

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, James H. Martin, <i>Speech and Language Processing</i> , 3rd edition (online), 2018.
2.	Christopher D. Manning, Hinrich Schutze, <i>Foundations of Statistical Natural Language Processing</i> , MIT Press, 1999.
3.	Yoav Goldberg, <i>Neural Network Methods for Natural Language Processing</i> , Morgan & Claypool Publishers, 2017.
4.	Steven Bird, Evan Klein, Edward Loper, <i>Natural Language Processing with Python</i> , O'Reilly Media Inc., 2009.

2.2 Additional Readings

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