Academic Org: Div of Computer Science & Engg – Subject: Computer Science

Course: CSCI5640   **Course ID:** 013795   **Eff Date:** 2022-07-01   **Crse Status: Active**   **Apprv. Status:** Approved   【Course Rev】

Natural Language Processing 自然語言處理

Natural language processing (NLP) is a crucial part of artificial intelligence (AI), which aims to endow computers with the ability to process human language. This course gives an overview of modern deep learning techniques for natural language processing. The course starts with basic linguistic concepts in NLP and moves from shallow bag-of-words representations to richer structural embeddings, which is the foundation for the successful use of deep learning in NLP. Then the course will guide you through three fundamental tasks of NLP: language modeling (LM), natural language understanding (NLU), and natural language generation (NLG), followed by some recent advances such as BERT and adversarial learning. Along the way we will introduce cutting-edge computational models together with insights from a linguistic perspective.

Advisory: Students are expected to have the background of deep learning, machine learning, linear algebra, programming , and data structure.

自然語言處理（NLP）是人工智能（AI）的重要組成部分，旨在使計算機具備處理人類語言的能力。 本課程概述了用於自然語言處理的現代深度學習技術。 該課程從NLP中的基本語言概念開始，講述從簡單的單詞袋表示到更豐富的結構性詞嵌入，這是在NLP中成功使用深度學習的基礎。 然後，本課程將指導您認識NLP的三個基本任務：語言建模（LM），自然語言理解（NLU）和自然語言生成（NLG），最後會介紹關於BERT和對抗性學習等最新進展。 在此過程中，我們將介紹最先進的計算模型並且從語言學角度闡述其原理。

建議: 學生應曾修讀深度學習，機器學習，線性代數，編程和數據結構的相關課程。

Grade Descriptor:

A

EXCELLENT – exceptionally good performance and far exceeding expectation in all or most of the course learning outcomes; demonstration of superior understanding of the subject matter, the ability to analyze problems and apply extensive knowledge, and skillful use of concepts and materials to derive proper solutions.

有關等級說明的資料，請參閱英文版本。

B

GOOD – good performance in all course learning outcomes and exceeding expectation in some of them; demonstration of good understanding of the subject matter and the ability to use proper concepts and materials to solve most of the problems encountered.

有關等級說明的資料，請參閱英文版本。

C

CU_CURR501
Page 2 of 4

THE CHINESE UNIVERSITY OF HONG KONG
Print Course Catalog Details

May 10, 2023
16:28:10 PM

FAIR – adequate performance and meeting expectation in all course learning outcomes; demonstration of adequate understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

D

MARGINAL – performance barely meets the expectation in the essential course learning outcomes; demonstration of partial understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

F

FAILURE – performance does not meet the expectation in the essential course learning outcomes; demonstration of serious deficiencies and the need to retake the course.

有關等級說明的資料，請參閱英文版本。

Equivalent Offering:
Units: 3 (Min) / 3 (Max) / 3 (Acad Progress)
Grading Basis: Graded
Repeat for Credit: N
Multiple Enroll: N
Course Attributes: MSc Computer Science
MPhil-PhD Computer Sci & Erg

Topics:

COURSE OUTCOMES

Learning Outcomes:

At the end of the course of studies, students will have acquired the ability to
1. Understand basic concepts in NLP from both computational and linguistic perspectives

CU_CURR501
Page 3 of 4

THE CHINESE UNIVERSITY OF HONG KONG
Print Course Catalog Details

May 10, 2023
16:28:10 PM

2.Understand fundamental tasks in NLP and its representative applications
3.Hands-on techniques to preprocess and analyze texts
4. **Implement deep learning models to resolve some simple real-world applications, such as sentiment analysis for tweets and dialog agents**

Course Syllabus:

At the end of the course of studies, students will have acquired

1. Knowledge in basic concepts of Natural Language Processing and Deep Learning
2. Knowledge in word embedding models and algorithms
3. Knowledge in language modeling techniques and approaches
4. Knowledge in natural language understanding and generation models and algorithms
5. Knowledge in some applications of natural language processing

Assessment Type:

Essay test or exam	: 20%
Presentation	: 20%
Project	: 20%
Short answer test or exam	: 40%

Feedback for Evaluation:

1. Quiz and examinations
2. Course evaluation and questionnaire
3. Question-and-answer sessions during class
4. Student consultation during office hours or online

Required Readings:

To be provided by course instructor.

Recommended Readings:

1. Jacob Eisenstein, "Introduction to Natural Language Processing", The MIT Press, 2019

CU_CURR501
Page 4 of 4

THE CHINESE UNIVERSITY OF HONG KONG
Print Course Catalog Details

May 10, 2023
16:28:10 PM

2. Dan Jurafsky and James H. Martin, "Speech and Language Processing, 2nd Edition", Prentice Hall, 2009. Third edition draft is available at web.stanford.edu/~jurafsky/slp3/.

OFFERINGS

1. CSCI5640   Acad Organization=CSEGV; Acad Career=RPG

COMPONENTS

LEC : Size=30; Final Exam=N; Contact=3
TUT : Size=30; Final Exam=N; Contact=1

ENROLMENT REQUIREMENTS

1. CSCI5640   **Enrollment Requirement Group:**
For students in MSc Computer Science; or
For students in MPhil-PhD Computer Science & Engineering; or
For undergraduate students in Computer Science (CSCIU & CSCIN) or Computer Engineering (CENGU & CENGN)

CAF

< E N D O F R E P O R T >