

NATURAL LANGUAGE PROCESSING

Version: 2024.1

1. THÔNG TIN CHUNG
GENERAL INFORMATION

Tên học phần	Natural language processing
Course name:	
Mã học phần	IT4772E
Code:	
Khối lượng	3(3-1-0-6)
Credit:	<ul style="list-style-type: none">- Lý thuyết - Lecture: 45 hours- Bài tập lớn - Midterm project: 15 hours- Thí nghiệm - Experiments: 0 hours
Học phần tiên quyết	No
Prerequisite:	
Học phần học trước	<ul style="list-style-type: none">- IT3010E: Data structures and algorithms
Prior course:	<ul style="list-style-type: none">- IT3190E: Machine learning- IT3320E: Introduction to deep learning
Học phần song hành	No
Paralell course:	

2. MÔ TẢ HỌC PHẦN - COURSE DESCRIPTION

This course presents fundamental problems and techniques in NLP. The contents include: core problems such as word formation, syntactic parsing, semantic parsing; information extraction such as named entity recognition; applications such as question answering and machine translation; speech processing such as speech recognition and speech synthesis. After the course, students could understand problems and apply techniques in NLP.

3. MỤC TIÊU VÀ CHUẨN ĐẦU RA CỦA HỌC PHẦN

GOAL AND OUTPUT REQUIREMENT

Sinh viên hoàn thành học phần này có khả năng
After this course the student will obtain the followings:

Mục tiêu/CDR Goal	Mô tả mục tiêu/Chuẩn đầu ra của học phần Description of the goal or output requirement	CDR được phân bổ cho HP/ Mức độ (I/T/U) Output division/ Level (I/T/U)
[1]	[2]	[3]
M1	Understand and apply technique to solve core NLP problems	1.1.2;1.2.1;1.2.5
M2	Understand and apply techniques to information	1.1.2;1.2.1;1.2.5

Mục tiêu/CĐR Goal	Mô tả mục tiêu/Chuẩn đầu ra của học phần Description of the goal or output requirement	CĐR được phân bổ cho HP/ Mức độ (I/T/U) Output division/ Level (I/T/U)
	extraction	
M3	Understand and apply techniques in NLP applications	1.1.2;1.2.1;1.2.5
M4	Understand and apply techniques in speech processing	1.1.2;1.2.1;1.2.5

4. TÀI LIỆU HỌC TẬP

Reference

Textbook

Reference book

- [1] Daniel Jurafsky and James H. Martin (2000, 2018). *Speech and Language Processing*. 2nd edition, 3rd edition draft. Prentice Hall
- [2] Christopher D. Manning and Hinrich Schütze (1999). *Foundations of Statistical Natural Language Processing*. MIT Press.
- [3] Yoav Goldberg (2017). *Neural Network Methods in Natural Language Processing*. Morgan & Claypool Publishers

5. CÁCH ĐÁNH GIÁ HỌC PHẦN - EVALUATION

Điểm thành phần Module	Phương pháp đánh giá cụ thể Evaluation method	Mô tả Detail	CĐR được đánh giá Output	Tỷ trọng Percent
[1]	[2]	[3]	[4]	[5]
A1. Điểm quá trình Mid-term (*)	Đánh giá quá trình Progress			40%
	A1.1. Continuous assessments	Presentation	M1.3; M3.1; M3.2	20%
	A1.2. Group project (assignment)			20%
A2. Điểm cuối kỳ Final term	A2.1. Thi cuối kỳ Final exam	Multi choice questions	M1÷M2	60%

* Điểm quá trình sẽ được điều chỉnh bằng cách cộng thêm điểm chuyên cần. Điểm chuyên cần có giá trị từ -2 đến +1, theo Quy chế Đào tạo đại học hệ chính quy của Trường ĐH Bách khoa Hà Nội.

The evaluation about the progress can be adjusted with some bonus. The bonus should belong to [-2, +1], according to the policy of Hanoi University of Science and Technology.

6. KẾ HOẠCH GIẢNG DẠY - SCHEDULE

Tuần Week	Nội dung Content	CDR học phần Output	Hoạt động dạy và học Teaching activities	Bài đánh giá
[1]	[2]	[3]	[4]	[5]
1	Overview of natural language processing Language modelling	M1, M2, M3	Note reading; Teaching;	A2.1
2	Word segmentation	M1	Note reading; Teaching;	A1.1 A2.1
3	Part-of-speech tagging	M1	Note reading; Teaching;	A1.1 A2.1
4	Constitutional parsing	M1	Note reading; Teaching;	A1.1 A2.1
5	Constitutional parsing	M1	Note reading; Teaching;	A1.1 A2.1
6	Probabilistic parsing	M1	Note reading; Teaching;	A1.1 A2.1
7	Dependency parsing	M1	Note reading; Teaching;	A1.1 A2.1
8	Semantic_parsing	M1	Note reading; Teaching;	A1.1 A2.1
9	Progress report	M3	Team work; Presentation;	A1.2
10	Lexical Semantics	M2	Note reading; Teaching;	A1.1 A2.1
11	Information Extraction	M3	Note reading; Teaching;	A1.1 A2.1
12	Question answering	M3	Note reading; Teaching;	A1.1 A2.1
13	Machine translation	M3	Note reading; Teaching;	A1.1 A2.1
14	Machine translation (con't)	M3	Note reading; Teaching;	A1.1 A2.1
15	Midterm project evaluation	M4	Note reading; Teaching;	A1.2
16	Midterm project evaluation	M4	Note reading; Teaching;	A1.2
17	Midterm project evaluation	M4	Note reading; Teaching;	A1.2

7. QUY ĐỊNH CỦA HỌC PHẦN - COURSE REQUIREMENT

(The specific requirements if any)

8. NGÀY PHÊ DUYỆT - DATE:

Chủ tịch hội đồng
Committee chair

Nhóm xây dựng đề cương
Course preparation group

Nguyễn Kiêm Hiếu
Lê Thanh Hương
Nguyễn Thị Thu Trang

9. QUÁ TRÌNH CẬP NHẬT - UPDATE INFORMATION

STT No	Nội dung điều chỉnh Content of the update	Ngày tháng được phê duyet Date accepted	Áp dụng từ kỳ/ khóa A pplicable from	Ghi chú Note
1			
2			