

ĐỀ CƯƠNG MÔN HỌC

CS524 – Một số ứng dụng của Xử lý ngôn ngữ tự nhiên

1. THÔNG TIN CHUNG

Tên môn học (tiếng Việt):	Một số ứng dụng của Xử lý ngôn ngữ tự nhiên
Tên môn học (tiếng Anh):	Applied Natural Language Processing
Mã môn học:	CS524
Thuộc khối kiến thức:	Tự chọn
Khoa/Bộ môn phụ trách:	Bộ môn Xử lý ngôn ngữ tự nhiên
Website môn học	
Giảng viên phụ trách:	ThS. Nguyễn Văn Toàn Email: toannv@uit.edu.vn
Giảng viên tham gia giảng dạy:	Các giảng viên của bộ môn
Số tín chỉ:	4 TC lý thuyết : 3 TC thực hành : 1
Lý thuyết: (tiết)	45
Thực hành: (tiết)	30
Tự học: (tiết)
Tính chất của môn	Tự chọn
Điều kiện đăng ký: (<i>môn học tiên quyết, học trước, song hành...</i>)	Xử lý ngôn ngữ tự nhiên

2. MỤC TIÊU MÔN HỌC

Môn học sẽ giới thiệu một số ứng dụng của xử lý ngôn ngữ tự nhiên trên máy tính được ứng dụng trong thực tế: xử lý dữ liệu, khai thác thông tin, tìm kiếm thông tin, phân tích tâm lý

Ngoài ra, môn học này cũng giới thiệu một số công cụ dùng để xử lý ngôn ngữ tự nhiên trên máy tính.

3. TÓM TẮT NỘI DUNG MÔN HỌC

Môn học cung cấp cho sinh viên một số mô hình ứng dụng, một số công cụ của xử lý ngôn ngữ tự nhiên trên máy tính.

4. CHUẨN ĐẦU RA

Mã số	Chuẩn đầu ra của môn học
[1] Kiến thức	
L1.	Kiến thức cơ bản về cách vận dụng các kỹ thuật của XLNNTN vào việc giải quyết các vấn đề của cuộc sống
L2.	Sinh viên nắm được cấu trúc, cách sử dụng một số các bộ ngữ liệu thông dụng
L3.	Sinh viên nắm được một số kỹ thuật máy học thông dụng trong xử lý ngôn ngữ tự nhiên.
L4.	Sinh viên nắm được nguyên lý hoạt động của một số ứng dụng/hệ thống trong xử lý ngôn ngữ tự nhiên.
[2] Kỹ năng	
	<p>Làm việc nhóm, tìm kiếm, đánh giá</p> <p>Phân tích, giải quyết vấn đề:</p> <ul style="list-style-type: none"> - Xác định, nêu vấn đề, mô hình hóa - Đánh giá giải pháp, nêu kiến nghị <p>Thử nghiệm, nghiên cứu, khám phá tri thức:</p> <ul style="list-style-type: none"> - Đề xuất giả thuyết - Khảo sát, thực nghiệm, kiểm chứng và bảo vệ giả thuyết
[3] Thái độ	
	<p>Tính cách cá nhân: Sáng tạo</p> <p>Yêu quê hương đất nước, ngôn ngữ tiếng mẹ đẻ</p>

5. NỘI DUNG CHI TIẾT

Tuần/ Thời lượng	Nội dung	Ghi chú/Mô tả hoạt động	Chuẩn đầu ra
	Chương 1. Giới thiệu 1.1. Xử lý ngôn ngữ tự nhiên 1.2. Một số vấn đề của XLNNTN 1.3. Ứng dụng trong XLNNTN 1.3. Một số công cụ của XLNNTN		L1, L2, L3
	Chương 2. Một số bộ ngữ liệu thông dụng 2.1. Phân loại 2.2. PennTree Bank 2.3. Google N-Grams Corpus 2.4. Brown Corpus 2.5. VLSP		L2
	Chương 3. Máy học trong XLNNTN 2.1. Một số kỹ thuật, mô hình máy học trong XLNNTN: thống kê, Bayes, EM, TBL, SVM 2.2. Một số ứng dụng		L3, L4
	Chương 4. Một số công cụ XLNNTN 4.1. Weka 4.2. Early 4.3. Giza++ 4.4. MOSE		L3, L4
	Chương 5: Một số ứng dụng XLNNTN thông dụng - Xây dựng ngữ liệu - Rút trích thông tin từ internet - Bắt lỗi chính tả (SP) - Kỹ thuật điền dấu tiếng Việt tự động - Ứng dụng XLNNTN trong học tiếng Anh - Xây dựng từ điển, ontology - Dịch máy (MT)		L2, L3, L4
	Chương 6. Một số ứng dụng xử lý âm thanh 3.1. Tổng quan 3.2. TTS 3.3. STT		L3, L4

6. PHƯƠNG PHÁP GIẢNG DẠY VÀ HỌC TẬP

- ❖ Phương pháp giảng dạy
 - Giáo viên giới thiệu, lý thuyết, tổng quan
 - Giáo viên thực hiện một case study làm ví dụ
 - Sinh viên thực hiện một đồ án nhỏ
- ❖ Đánh giá:

- Đồ án môn học.

7. HÌNH THỨC ĐÁNH GIÁ KẾT QUẢ HỌC TẬP

❖ Hình thức đánh giá:

- Đồ án nhóm: 100%

8. TÀI LIỆU HỌC TẬP, THAM KHẢO

Sách, giáo trình chính

Giáo trình Xử Lý Ngôn Ngữ Tự Nhiên, First Edition, Đinh Điền, NXB ĐHQG-HCM, 2006.

Sách tham khảo

Machine Learning for Natural Language Processing, Martin Emms-mtemms@cs.tcd.ie, Saturnino Luz,- luzs@cs.tcd.ie; January 13, 2011

Trưởng khoa/ bộ môn

(Ký và ghi rõ họ tên)

Giảng viên

(Ký và ghi rõ họ tên)

Nguyễn Văn Toàn