Trường Đại học Bách Khoa - ĐHQG Tp.HCM
Khoa: Khoa Khoa học và Kỹ thuật Máy tính
Khoa/Bộ môn quản lý MH: Khoa học Máy tính

Tp.HCM, ngày ..... tháng ..... năm .....

Đề cương môn học Sau đại học

CÁCH TIẾP CẬN HIỆN ĐẠI TRONG XỬ LÝ TIẾNG NÓI (MODERN APPROACHES FOR SPEECH PROCESSING)

Mã số MH: CO5257

Số tín chỉ:	**Tc (LT.BT&TH.Tự Học): 3**						**ECTS: 6**				
Số tiết -Tổng:	**75**	LT:	**30**	BT:	**0**	TH:	**0**	ĐA:		BTL/TL:	**45**
Đánh giá:	Tiểu luận:			50%							
	Thi:			50%							
- Môn tiên quyết:											
- Môn học trước:	- Trí tuệ nhận tạo									MS: 054010	
- Môn song hành:	- Học máy và Ứng dụng									MS: 055241	
- CTĐT ngành (Mã ngành):	(848010), Khoa Học Máy Tính (8480101)										
- Ghi chú khác:											

1. <u>Mục tiêu môn học:</u>

- Cung cấp cho học viên các kiến thức cơ bản về xử lý giọng nói

- Giới thiệu các cách tiếp cận trong nhận dạng giọng nói

- Giới thiệu cách mô hình học sâu trong nhận dạng giọng nói

- Giới thiệu hệ thống sinh giọng nói từ văn bản

Aims:

2. <u>Nội dung tóm tắt môn học:</u>

Môn học giới thiệu các khái niệm trong xử lý giọng nói, từ cơ bản đến các cách tiếp cận nâng cao. Nội dung môn học gồm các chủ đề sau:

- Giới thiệu lịch sử của xử lý giọng nói và các kiến thức cơ bản về âm vị

- Các cách tiếp cận trong nhận dạng giọng nói.

- Các mô hình học sâu nâng cao cho nhận dạng giọng nói

- Tác nhân đối thoại

- Hệ thống sinh giọng nói từ văn bản

Course outline:

The course introduces concepts in speech processing, from basic to advanced approaches. The content of the course is given as follows:

- Introduce the history of speech processing and basic knowledge on phonetics

- Approaches in speech recognition

- Advanced deep models for speech recognition

- Conversational agent

- Text to speech system

3. Tài liệu học tập:

[1] Quan Thanh Tho, Advanced Deep Learning Models, VNUHCM Press (in Vietnamese)

[2] Dan Jurafsky and James H. Martin. *Speech and Language Processing (3rd)*

[3] Yoav Goldberg. *A Primer on Neural Network Models for Natural Language Processing*

[4] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press.

4. Các hiểu biết, các kỹ năng cần đạt được sau khi học môn học:

STT	Chuẩn đầu ra môn học (CĐRMH)	Công cụ đánh giá CĐRMH	Đóng góp CĐR Chương trình (CĐRCT)		
			Ứng dụng	Nghiên cứu	

Learning outcomes:

No.	Course learning outcomes (CLO)	CLO assessment	Matching with PLO		
			Coursework	Research	

Bảng ánh xạ chuẩn đầu ra môn học và chuẩn đầu ra chương trình ứng dụng:

	Chuẩn đầu ra của chương trình (CĐRCT)										
Chuẩn đầu ra môn học (CĐRMH)	a	b	c	d	e	f	g	h	i	j	k

Bảng ánh xạ chuẩn đầu ra môn học và chuẩn đầu ra chương trình nghiên cứu:

	Chuẩn đầu ra của chương trình (CĐRCT)										
Chuẩn đầu ra môn học (CĐRMH)	a	b	c	d	e	f	g	h	i	j	k

5. Hướng dẫn cách học - chi tiết cách đánh giá môn học:

- Dự các buổi trên lớp và tích cực tham gia các hoạt động thảo luận trên lớp.

- Thực hành các giải thuật, mô hình đã được giới thiệu ở nhà

- Chủ động trong quá trình làm tiểu luận.

Learning strategies & Assessment Scheme:

- Attend the lectures and be active in discussion activities

- Practice algorithms, models at home

- Be active while conducting the course project

6. Nội dung chi tiết:

Tuần/ Buổi	Chủ đề (chương)	Nội dung	Chuẩn đầu ra môn học	Tài liệu
1	Tổng quan môn học, lịch sử, và khái niệm ngữ âm	Giới thiệu về các nội dung môn học, lịch sử của xử lý giọng nói, và các khái niệm cơ bản về ngữ âm		
2	Nhận dạng giọng nói: Mô hình nhiễu kênh, HMMs, Forward, Viterbi, Word Error Rate	Giới thiệu các cách tiếp cận cơ bản trong nhận dạng giọng nói và trình bày các giải thuật truyền thống.		
3	Nhận dạng giọng nói: giải mã nâng cao, đầu dò trạng thái hữu hạn	Giới thiệu các cách tiếp cận nâng cao trong xử lý giọng nói		
4	Nhận dạng giọng nói: mô hình âm học GMM và trích xuất đặc trưng	Giới thiệu mô hình âm học GMM và trích xuất đặc trưng trong xử lý giọng nói.		
5	Nhận dạng giọng nói: những mô hình âm học sử dụng mạng neural	Giới thiệu các mô hình mạng neural sử dụng trong trích xuất đặc trưng âm học		
6	Nhận dạng giọng nói: mô hình nhận dạng mạng neural end-to-end	Giới thiệu mạng neural end-to-end dùng trong nhận dạng giọng nói		
7	Các tác nhân đối thoại	Giới thiệu các hệ thống hội thoại: frame-based, Dialog Acts, Information State, Markov Decision Processes, và các cách tiếp cận học sâu		
8	Text to Speech (TTS): □tổng quan, chuẩn hoá văn bản, □từ chữ sang âm thanh, tiến trình	Giới thiệu tổng quan về TTS, các thành phần cơ bản của quy trình		
9	Text to Speech (TTS): mô hình gắn kết, mô hình tham số, và mô hình End-to End với mạng Neural	Giới thiệu các mô hình dùng trong TTS		
10	Những vấn đề mở rộng trong TTS	Giới thiệu các vấn đề liên quan đến hệ thống TTS		

7. Giảng viên tham gia giảng dạy:

CBGD chính:	TS. Nguyễn Đức Dũng

CBGD tham gia:	PGS.TS Quản Thành Thơ

XÁC NHẬN CỦA HỘI ĐỒNG XÂY DỰNG CHƯƠNG TRÌNH ĐÀO TẠO VÀ KHOA

Tp. Hồ Chí Minh, ngày ..... tháng ..... năm .....

GIẢNG VIÊN LẬP ĐỀ CƯƠNG

TS. Phạm Hoàng Anh