

확장강의계획서

(2022 년도 2 학기)

과목명	음성인식 (2022년 1학기까지 개설되었던 대화형사용자인터페이스개론과 동일한 코드를 사용하는 과목으로, 대화형사용자인터페이스개론을 수강한 학생이 수강하는 경우 재수강이 됩니다)	과목번호	CESG311/CSE5311
구분(학점)	3학점	수강대상	석사과정/학부4학년
수업시간	화/목 4:30~5:45PM	강의실	

담당교수 (사진)	성명: 김지환	과목홈페이지: eclass.sogang.ac.kr
	E-mail: kimjihwan@sogang.ac.kr	연락처: 이메일
	장소: AS 713 면담시간: 이메일로 사전 예약	

I. 교과목 개요(Course Overview)

1. 수업개요							
본과목의 목표는 대화형 사용자 인터페이스 시스템을 구성하고 있는 주요 요소인 음성인식의 기본 개념을 이해하는 것이다. 딥러닝 기반의 음성인식 주요 기술 및 최근 기술 동향을 DNN-WFST, CTC, RNN-T, Seq2Seq, Transformer 중심으로 살펴보고, 음향모델은 DNN-HMM 을 중심으로 언어모델은 n-gram 중심으로, 디코딩 네트워크는 WFST 중심으로 공부한다. 그리고, 음성인식 관련 오픈 소스 기반 프로그램을 소개한다.							
2. 선수학습내용							
없음							
3. 수업방법 (%)							
강의	토의/토론	실험/실습	현장학습	개별/팀 별 발표	기타		
100 %	%	%	%	%	%		
4. 평가방법 (%)							
중간고사	기말고사	퀴즈	발표	프로젝트	과제물	참여도	기타
50 %	50 %	%	%	%	%	%	%

II. 교과목표(Course Objectives)

본과목의 목표는 대화형 사용자 인터페이스 시스템을 구성하고 있는 주요 요소인 음성인식 의 기본 개념을 이해 하는 것이다. 딥러닝 기반의 음성인식 주요 기술 및 최근 기술 동향을 DNN-WFST, CTC, RNN-T, Seq2Seq, Transformer 중심으로 살펴본다. 음향모델은 DNN-HMM을 중심으로, 언어모델은 n-gram 중심으로, 디코딩 네트워크는 WFST 중심으로 이해한다

III. 수업운영방식(Course Format)

(* 1-3의 수업방법의 구체적 설명)

학교의 방역지침에 따름 (2022년 7월말 기준 대면 수업으로 진행)

IV. 학습 및 평가활동(Course Requirements and Grading Criteria)

-본 과목은 중간고사 (50%), 기말고사 (50%)로 평가한다.

1 중간고사

- 강의 시간에 배운 핵심 알고리즘의 이해도를 측정한다

2 기말고사

- 강의 시간에 배운 핵심 알고리즘의 이해도를 측정한다

V. 수업규정(Course Policies)

출석은 성적에 반영하지 않음. 비대면 강의에 대한 학교의 FA 방침에 따라 FA는 적용함

VI. 교재 및 참고문헌(Materials and References)

주교재: 강의 노트

참고자료

1. Daniel Jurafsky and James H. Martin : Speech and Language Processing, Pearson

International Edition, Second Edition, 2009
2. Xuedong Huang, Alex Accero, Hsiao Wuen Hon, Spoken Language Processing: A Guide to Theory, Algorithm and System Development, 2001
3. 관련 Paper, 수업 시간에 소개

VII. 주차별 강의계획(Course Schedule)

(* 추후 변경될 수 있음)

1 주차	학습목표	오리엔테이션 및 음성인식의 전반적인 개요에 대해 공부한다.
	주요학습내용	OT & 음성대화 인터페이스의 전반적인 개념
	수업방법	강의
	수업자료	강의노트
	과제	
2 주차	학습목표	딥러닝 기반의 음성인식 주요 기술 및 최근 기술 동향
	주요학습내용	DNN-WFST, CTC, RNN-T, Seq2Seq, Transformer
	수업방법	강의
	수업자료	강의노트
	과제	
3 주차	학습목표	음향모델과 HMM 개요
	주요학습내용	음향모델이 가져야 할 특성과 HMM에 대해 공부한다
	수업방법	강의
	수업자료	강의노트
	과제	
4 주차	학습목표	Forward Algorithm과 Viterbi Algorithm에 대해 공부한다.
	주요학습내용	Forward Algorithm & Viterbi Algorithm
	수업방법	강의
	수업자료	강의노트
	과제	
5 주차	학습목표	DNN 음향모델과 Tree-based state clustering에 대해 공부한다
	주요학습내용	DNN 음향모델 및 Tree-clustering

	수업방법	강의
	수업자료	강의노트
	과제	
6 주차	학습목표	Multi-layer-perceptron과 Back-propagation 문제를 직접 풀어본다
	주요학습내용	Multi-layer-perceptron 및 Back-propagation
	수업방법	강의
	수업자료	강의노트
	과제	
7 주차	학습목표	언어모델의 개요와 N-gram 언어모델에 대해 공부한다.
	주요학습내용	언어모델개요 및 N-gram 언어모델
	수업방법	강의
	수업자료	강의노트
	과제	
8 주차	학습목표	
	주요학습내용	중간고사
	수업방법	
	수업자료	
	과제	
9 주차	학습목표	Back-off 언어모델과 SRILM에서의 구현방식에 대해 공부한다.
	주요학습내용	Katz Back-off 및 구현 이슈사항
	수업방법	강의
	수업자료	강의노트
	과제	
10 주차	학습목표	Lexical tree 기반 Decoding network에 대해 공부한다
	주요학습내용	Lexical tree

	수업방법	강의
	수업자료	강의노트
	과제	
11 주차	학습목표	WFST 개요 및 G transducer에 대해 공부한다.
	주요학습내용	WFST 개요 및 G transducer
	수업방법	강의
	수업자료	강의노트
	과제	
12 주차	학습목표	예제를 통해 Arpa 포맷으로부터 G transducer를 생성 해 본다.
	주요학습내용	G transducer 생성
	수업방법	강의
	수업자료	강의노트
	과제	
13 주차	학습목표	L transducer와 C transducer에 대해 공부한다.
	주요학습내용	L transducer 및 C transducer
	수업방법	강의
	수업자료	강의노트
	과제	
14 주차	학습목표	H transducer와 WFST 생성에 필요한 연산에 대해 공부한다.
	주요학습내용	H transducer 및 WFST 연산
	수업방법	강의
	수업자료	강의노트
	과제	

15 주차	학습목표	WFST 연산 및 음성인식 toolkit들을 소개한다.
	주요학습내용	WFST 연산과 음성인식 toolkit 소개
	수업방법	강의
	수업자료	강의노트
	과제	
16 주차	학습목표	
	주요학습내용	기말고사
	수업방법	
	수업자료	
	과제	

VIII. 참고사항(Special Accommodations)

--