جـــامـــعـــة مـحـمـد بـــن زايـــد
للـــذكـــاء الاصـــطـــنـــاعـــي
MOHAMED BIN ZAYED UNIVERSITY
OF ARTIFICIAL INTELLIGENCE

Core Courses Syllabi

NLP703 - Speech Processing

Title	Speech Processing
Code	NLP703
Loading	4 Credit-hours
Prerequisites	• Basic Concepts in Linear Algebra, Calculus, Probability and Statistics • Programming in Python or similar language
Catalog Description	This course provides a comprehensive introduction to Speech Processing. It builds upon fundamental concepts in Speech Processing and assumes familiarization with Mathematical and Signal Processing concepts.
Goal	This graduate level course aims to equip students with deep understanding of foundations of core speech processing algorithms.
Content	The course covers the following major modules: **(I)** Speech Recognition, **(II)** Speech Synthesis and **(III)** Dialogue and Conversational System
Recommended Textbooks	1. Chris Manning et al, *Foundation of statistical natural language processing*, MIT Press (1999) ISBN: 0262133601 2. Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press. ISBN:
Recommended References & Supplemental Material	Relevant research papers, technical reports, and surveys for each topic, where needed, are identified in the teaching plan ahead. In addition, the following textbook may be useful: C. Bishop, *Pattern Recognition and Machine Learning*, Berlin: Springer-Verlag, 2006. ISBN: 0387310738

Teaching Week	Topics
1	**Course Overview** **Lecture** • Introduction to speech processing, digitization and recording of the speech signal. • Overview of speech processing applications **Lab** • Demonstration of the development platform and libraries used for speech and language processing (e.g. Tensorflow, PyCharm, Pure Data) • Warm-up exercise for Python and Jupyter Notebook • Start **Project**
2	**Speech Processing: Basics** **Lecture** • Acoustic phonetics and articulatory phonetics (The International Phonetic Alphabet, IPA) • The syllable • Consonants and vowels • Phonetic transcription • Machine-readable phonetic alphabets **Lab** • Assignment 1
3	**Speech Processing: Human Speech Production System** **Lecture** • Introduction of the human speaking system, including the vocal tract, the palate and lips etc. • Introduce Uniform Tube Modeling in speech production • Introduce voiced and unvoiced segments, first and second formant **Lab** • Assignment 1
4	**Speech Processing: Human Speech Perception System** **Lecture** • Introduction of the basic mechanisms of human hearing, including the ear canal, eardrums and pinna, etc • Description of action of the cochlea **Lab** • Basic programming exercises for using tools for speech synthesis (e.g. implementing the Holmes parallel formant synthesizer)

Teaching Week	Topics
5	**Speech Signal Analysis: Forms of the Speech Signal** **Lecture** • Describe the different forms of signals • Discrete/continuous signal • Signal Processing • Sensors/transducers • Microphones • Periodic signals (sines & cosines) • White noise • Stationarity **Lab** • Assignment 2
6	**Speech Signal Analysis: Feature Representation** **Lecture** • Overview of techniques for representing speech Signal as features • Feature extraction, extraction of the fundamental frequency • Harmonics • The spectrum • The complex spectrum • Impulse response • The source-filter model of speech **Lab** • Programming exercise for feature extraction from the speech signal • Assignment 2
7	**Speech Signal Analysis: Models** **Lecture** • Overview of different techniques used for speech signal analysis • Waveform analysis • Spectrographic analysis • Binaural processing • Fourier Transform **Lab** • Programming exercises for implementation of Fourier Transform

Teaching Week	Topics
8	**Revision** **Lecture and Lab** • Project Work • Assignment 2
9	**Speech recognition: Basics** **Lecture** • Introduction of basic speech recognition models • Noisy Channel Model • Hidden Markov Models • Word error rate (evaluation method) **Lab** • Programming exercises for implementation of Hidden Markov Model
10	**Speech Recognition: Advanced** **Lecture** • Introduction to advanced speech recognition models • Forward and backward Viterbi • Advanced decoding • Finite State Transducers • Gaussian mixture model **Lab** • Programming exercises for implementation of the Forward and backward Viterbi decoding algorithm
11	**Neural Speech Recognition: Basics** **Lecture** • Introduction to neural network acoustic modellings • Multi-layer Feedforward Neural Network • Recurrent Neural Networks (RNNs) • Convolutional Neural Networks **Lab** • Programming exercises for implementation of MLP

Teaching Week	Topics
12	**Neural Speech Recognition: End-to-end Neural Network Speech Recognition** **Lecture** • Description of the processing pipeline used in practical speech recognition systems, covering neural models used in individual modules in the pipeline, from signal extraction to speech recognition **Lab** • Assignment 3
13	**Conversational Agents System: Introduction and Frame-based Dialogue** **Lecture** • Overview of different conversational systems (e.g. goal oriented v.s conversation systems) • Description of the techniques used for frame extraction and recognition in corpora • Semantic Parsing • Conditional Random Field **Lab** • Programming exercises for implementation of semantic parser
14	**Conversational Agents System: Utterance Understanding, Dialogue Management and Response Generation** **Lecture** • Models for dialog act classification (e.g. statement, rephrase, questions, etc.) • Frameworks for dialogue management • Markov Decision Processes • Information state • Reinforcement learning **Lab** • Programming exercises for implementation of dialog act classifier
15	**Conversational Agents System: Neural Approaches used in Conversational Agents System** **Lecture** • Introduction to neural networks models used in conversational agents systems • Sequence to sequence neural models **Lab** • Review and exam preparation • Assignment 4