SPEECH PROCESSING

[Revised Credit System]

(Effective from the academic year 2022-23)

Program Elective

[Revised Credit System]

Subject Code	**CSE 4430 [3 0 0 3]**	**IA Marks**	**50**
Number of Lecture Hours/Week	**03**	**Exam Marks**	**50**
Total Number of Lecture Hours	**36**	**Exam Hours**	**03**

CREDITS – 03

Course objectives: This course will enable students to

- Understand the production of human sound and various techniques involved in speech analysis.
- Comprehend the modeling of speech using various techniques.
- Analyze large vocabulary for continuous speech recognition.
- Describe the techniques involved in speech synthesis and assessment of its quality.

Module -1	**Teaching Hours**
BASIC CONCEPTS Speech Fundamentals: Articulatory Phonetics – Production and Classification of Speech Sounds; Acoustic Phonetics – acoustics of speech production; Review of Digital Signal Processing concepts; Short-Time Fourier Transform, Filter-Bank and LPC Methods **Text Book 1:** Chapter 1,2,3 [1.1, 2.1, 2.2, 2.3, 2.4, 2.5, 3.2, 3.3]	**8 Hours**
Module -2	
SPEECH ANALYSIS Features, Feature Extraction and Pattern Comparison Techniques: Speech distortion measures – mathematical and perceptual – Log Spectral Distance, Cepstral Distances, Weighted Cepstral Distances and Filtering, Likelihood Distortions, Spectral Distortion	**8 Hours**

using a Warped Frequency Scale, LPC, PLP and MFCC Coefficients, Time Alignment and Normalization – Dynamic Time Warping, Multiple Time – Alignment Paths. **Text Book 1 :** Chapter 4 [4.1, 4.2, 4.3, 4.4, 4.5, 4.7]	
Module – 3	
SPEECH MODELING Hidden Markov Models: Markov Processes, HMMs – Evaluation, Optimal State Sequence – Viterbi Search, Baum-Welch Parameter Re-estimation, Implementation issues. **Text Book 2**: Chapter 6 [6.1,6.2, 6.3, 6.4, 6.5]	**7 Hours**
Module – 4	
SPEECH RECOGNITION Large Vocabulary Continuous Speech Recognition: Architecture of a large vocabulary continuous speech recognition system – acoustics and language models – n-grams, context dependent sub-word units; Applications and present status. **Text Book 1:** Chapter: 8	**8 Hours**
Module – 5	
SPEECH SYNTHESIS: Text-to-Speech Synthesis: Concatenative and waveform synthesis methods, sub-word units for TTS, intelligibility and naturalness – role of prosody, Applications and present status. **Text Book 2:** Chapter: 8 [8.4, 8.5, 8.6]	**5 Hours**
Course outcomes:	
After studying this course, students will be able to: 1. Comprehend the mechanism of production of human sound 2. Understand the various techniques for analysis of speech signals 3. Learn the mechanism of speech modeling 4. Understand continuous speech recognition 5. Specify the techniques involved in Speech synthesis	

Text Books:

1. Lawrence Rabinerand Biing-Hwang Juang, *“Fundamentals of Speech Recognition”*, Prentice Hall, 1993.

2. Daniel Jurafsky and James H Martin, *“Speech and Language Processing – An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition”*, 2nd edition, Pearson Education, 2008.

Reference Books:

1. Steven W. Smith, *“The Scientist and Engineer’s Guide to Digital Signal Processing”*, California Technical Publishing.
2. Thomas F Quatieri, *“Discrete-Time Speech Signal Processing – Principles and Practice”*, Pearson Education.
3. Claudio Becchetti and Lucio Prina Ricotti, *“Speech Recognition”, Theory and C++ implementation”*, John Wiley and Sons, 1999.
4. Ben gold and Nelson Morgan, *“Speech and audio signal processing”*, processing and perception of speech and music, Wiley- India Edition, 2006 Edition.
5. Frederick Jelinek, *“Statistical Methods of Speech Recognition”*, MIT Press.