

## SEMESTER S6

### SPEECH AND AUDIO PROCESSING

<b>Course Code</b>	<b>PEEVT 634</b>	<b>CIE Marks</b>	40
<b>Teaching Hours/Week (L: T:P: R)</b>	3:0:0:0	<b>ESE Marks</b>	60
<b>Credits</b>	3	<b>Exam Hours</b>	2 Hrs. 30 Min.
<b>Prerequisites (if any)</b>	None	<b>Course Type</b>	Theory

#### Course Objectives:

1. To impart the basic concepts of speech signal processing
2. To familiarize the auditory mechanism and speech perception

#### SYLLABUS

<b>Module No.</b>	<b>Syllabus Description</b>	<b>Contact Hours</b>
<b>1</b>	Speech Production: - Acoustic theory of speech production- -Source/Filter model - Pitch, Formant, Spectrogram-- Discrete model for speech production, Articulatory Phonetics -Acoustic Phonetics- Basic speech units and their classification.	<b>9</b>
<b>2</b>	Short-Time Speech Analysis, Windowing, STFT, spectra of windows- Wide and narrow band spectrogram -Time domain parameters (Short time energy, short time zero crossing Rate, ACF). Frequency domain parameters-Filter bank analysis. STFT Analysis. Prosody of speech. MFCC-computation, LPC Model, Pitch and Formant Estimation.	<b>9</b>
<b>3</b>	Speech Enhancement: Spectral subtraction and Filtering, Harmonic filtering, parametric resynthesis. Speaker Recognition: Speaker verification and speaker identification- log-likelihood. Machine learning models in Speaker Recognition. Language identification: implicit and explicit models.	<b>9</b>
<b>4</b>	Signal Processing models of audio perception: Basic anatomy of hearing System: Basilar membrane behaviour. Sound perception: Auditory Filter Banks, Critical Band Structure, Absolute Threshold of Hearing, Masking-	<b>9</b>

	Simultaneous Masking, Temporal Masking. Models of speech perception	
--	---------------------------------------------------------------------	--

**Course Assessment Method**  
(CIE: 40 marks, ESE: 60 marks)

**Continuous Internal Evaluation Marks (CIE):**

Attendance	Assignment/ Microproject	Internal Examination-1 (Written)	Internal Examination- 2 (Written )	Total
5	15	10	10	40

**End Semester Examination Marks (ESE)**

*In Part A, all questions need to be answered and in Part B, each student can choose any one full question out of two questions*

Part A	Part B	Total
<ul style="list-style-type: none"> <li>• 2 Questions from each module.</li> <li>• Total of 8 Questions, each carrying 3 marks</li> </ul> <p align="center"><b>(8x3 =24marks)</b></p>	<ul style="list-style-type: none"> <li>• Each question carries 9 marks.</li> <li>• Two questions will be given from each module, out of which 1 question should be answered.</li> <li>• Each question can have a maximum of 3 sub divisions.</li> </ul> <p align="center"><b>(4x9 = 36 marks)</b></p>	<b>60</b>

**Course Outcomes (COs)**

At the end of the course students should be able to:

Course Outcome		Bloom's Knowledge Level (KL)
<b>CO1</b>	To describe the fundamental concepts, principles, and theories of speech production	<b>K1</b>
<b>CO2</b>	To analyse the speech signal in the time and frequency domain	<b>K2</b>
<b>CO3</b>	To apply speech processing concepts in real-world applications	<b>K3</b>
<b>CO4</b>	To describe the fundamental concepts, principles, and theories of hearing mechanism	<b>K1</b>
<b>CO5</b>	To develop applications by combining concepts of speech production and hearing mechanism	<b>K3</b>

Note: K1- Remember, K2- Understand, K3- Apply, K4- Analyse, K5- Evaluate, K6- Create

### CO-PO Mapping Table (Mapping of Course Outcomes to Program Outcomes)

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3											
CO2	3	2										
CO3	3	2										
CO4	3											
CO5	3	2	3	3	3	3		2				

Note: 1: Slight (Low), 2: Moderate (Medium), 3: Substantial (High), -: No Correlation

Text Books				
Sl. No	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
1	Speech Communications: Human and Machine, 2nd Edition	Douglas O'Shaughnessy	Wiley-IEEE Press	2 <sup>nd</sup> edition
2	Discrete-Time Speech Signal Processing: Principles and Practice	Thomas F. Quatieri	Prentice-Hall Signal Processing Series	2001

Reference Books				
Sl. No	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
1	Digital Processing of Speech Signals	Rabinar	Pearson	2003

Video Links (NPTEL, SWAYAM...)	
Module No.	Link ID
1	Speech and Audio Processing 1: Introduction to Speech Processing - Professor E. Ambikairajah <a href="https://www.youtube.com/watch?v=Xjzm7S_kBU">https://www.youtube.com/watch?v=Xjzm7S_kBU</a>
2	Speech Analysis - Professor E. Ambikairajah <a href="https://www.youtube.com/watch?v=Y_mSQ7tTlvQ&amp;t=38s">https://www.youtube.com/watch?v=Y_mSQ7tTlvQ&amp;t=38s</a>
3	Speech and Audio Processing 1: Introduction to Speech Processing - Professor E. Ambikairajah <a href="https://www.youtube.com/watch?v=Xjzm7S_kBU">https://www.youtube.com/watch?v=Xjzm7S_kBU</a>
4	Video Links available on hearing anatomy