FACULTY OF ENGINEERING & TECHNOLOGY

Effective from Academic Batch: 2022-23

Programme: Bachelor of Technology (Artificial Intelligence (AI) and Data Science)

Semester: VIII

Course Code: 202047813

Course Title: Speech and Audio Signal Processing

Course Group: Professional Elective Course-V

Course Objectives: This course provides the basic knowledge regarding the mechanism of speech production, speech analysis and parametric representation of speech. It also provides an overview of various applications of speech. It also emphasis on Perception of Sound, Psycho-acoustic analysis, Spatial Audio Perception and rendering. Various speech compression schemes were also incorporated.

Teaching & Examination Scheme:

Contact hours per week			Course Credits	Examination Marks (Maximum / Passing)				
				Theory		J/V/P*		Total
Lecture	Tutorial	Practical		Internal	External	Internal	External	
3	0	2	4	50/ 18	50/ 17	25/ 09	25/ 09	150 / 53

* **J**: Jury; **V**: Viva; **P**: Practical

Detailed Syllabus:

Sr.	Contents	Hours
1	Speech Production: Acoustic theory of speech production. Speech Analysis: Short-Time Speech Analysis, Time domain analysis (Short time energy, short time zero crossing Rate, ACF). Parametric representation of speech: AR Model, ARMA model. LPC Analysis (LPC model, Auto correlation method).	07
2	Frequency domain analysis (Filter Banks, STFT, Spectrogram), Cepstral Analysis, MFCC. Fundamentals of Speech recognition and Text-to-speech. Speech coding, speech enhancement, Speaker Verification, Language Identification	08
3	Signal Processing Models of Audio Perception: Basic anatomy of hearing System. Auditory Filter Banks, Psycho-acoustic analysis: Critical Band Structure, Absolute Threshold of Hearing, Simultaneous Masking, Temporal Masking, Quantization Noise Shaping, MPEG psycho-acoustic model.	08
4	Audio compression methods: Sampling rate and bandwidth requirement for digital audio, Redundancy removal and Perceptual irrelevancy removal, Transform coding of digital audio: MPEG2-AAC coding standard, MDCT and its properties, Pre-echo and pre-echo suppression, Loss less coding methods.	08
5	Spatial Audio Perception and rendering: The physical and psycho-acoustical basis of sound localization and space Perception. Spatial audio standards. Audio quality analysis: Objective analysis methods- PEAQ, Subjective analysis methods - MOS score,	08

	MUSHRA score	

Suggested Specification table with Marks (Theory) (Revised Bloom's Taxonomy):

Distribution of Theory Marks						**R**: Remembering; **U**: Understanding; **A**: Application, **N**: Analyze; **E**: Evaluate; **C**: Create
R	**U**	**A**	**N**	**E**	**C**	
15%	20%	30%	20%	10%	--	

Note: This specification table shall be treated as a general guideline for students and teachers. The actual distribution of marks in the question paper may vary slightly from above table.

Reference Books:	
1	Douglas O'Shaughnessy, Speech Communications: Human & Machine, IEEE Press, Hardcover 2/e, 1999; ISBN: 0780334493
2	Nelson Morgan and Ben Gold, Speech and Audio Signal Processing: Processing and Perception Speech and Music, July 1999, John Wiley & Sons, ISBN: 0471351547
3	Donald G. Childers, Speech Processing and Synthesis Toolboxes, John Wiley & Sons, September 1999; ISBN: 0471349593
4	Rabiner and Juang, Fundamentals of Speech Recognition, Prentice Hall, 1994.
5	Rabiner and Schafer, Digital Processing of Speech Signals, Prentice Hall, 1978.
6	Thomas F. Quatieri, Discrete-Time Speech Signal Processing: Principles and Practice, Prentice Hall; ISBN: 013242942X; 1/e

Course Outcomes (CO):

Sr.	Course Outcome Statements	% weightage
CO-1	To familiarize the basic mechanism of speech production and the basic concepts of methods for speech analysis and parametric representation of speech.	**20**
CO-2	To give an overall picture about various applications of speech processing.	**30**
CO-3	To impart ideas of Perception of Sound, Psycho-acoustic analysis, Spatial Audio Perception and rendering.	**30**
CO-4	To introduce Audio Compression Schemes.	**20**

List of Practicals:

1	Write a program to read an Audio file and Play it.
2	Write a program to read an Audio file and play it by changing the sampling frequency.
3	Write a program to compute a Time domain parameter (Short time energy) of a speech signal.
4	Write a program to plot STFT (Short Time Fourier Transform).
5	Write a Program to implement Text to Speech conversion.
6	Write a program to perform Cepstral Analysis of speech by computing various signals.
7	Write a program to implement Simultaneous Masking.
8	Write a program to implement Temporal Masking.
9	Write a program to compress the given audio.
10	Write a program to perform quality analysis of a given audio.

Supplementary learning Material:

1. Coursera course on Speech and Audio signal processing
2. NPTEL course on Speech and Audio signal processing

Curriculum Revision:	
Version:	1.0
Drafted on (Month-Year):	June-2021
Last Reviewed on (Month-Year):	-
Next Review on (Month-Year):	June-2025