राष्ट्रीय प्रौद्योगिकी संस्थान पटना / NATIONAL INSTITUE OF TECHNOLOGY PATNA
संगणक विज्ञान एंव अभियांत्रिकी विभाग / DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
अशोक राजपथ, पटना-८००००५, बिहार / ASHOK RAJPATH, PATNA-800005, BIHAR
Phone No.: 0612-2372715, 2370419, 2370843, 2371929 Ext- 200, 202 Fax-0612-2670631 Website: www.nitp.ac.in

No:- **Date:**

CSX4182: Speech Processing
L-T-P-Cr.- 2-0-2-3

Prerequisites: (i) Digital Signal Processing and (ii) Probability, Random Variable and Stochastic Process

Objectives: This course concerns with analysis and processing of speech signals for development of different voice biometric based applications.

Course Outcome: Upon successful completion of this course, students should be able to:

Sl. No.	Course Outcome (CO)	Program Outcome (PO)
1	Explain the human speech production and perception mechanisms, including the role of different organs and the linguistic aspects of speech.	PO1
2	Analyze speech signals in the time and frequency domains using techniques like the Short-Term Fourier Transform (STFT) and filter banks.	PO1, PO2, PO5
3	Apply cepstral analysis techniques for speech feature extraction, including real, complex, and Mel-Frequency Cepstral Coefficients (MFCCs), along with delta and delta-delta features.	PO1, PO2, PO3, PO5
4	Implement Linear Prediction (LP) analysis for speech representation, utilizing Levinson-Durbin's method to obtain LP parameters like spectrum, cepstrum, and residual.	PO1, PO2, PO3, PO5
5	Describe the principles of sinusoidal analysis for speech and apply techniques like phase unwrapping for speech signal synthesis.	PO1, PO3
6	Design and evaluate speech processing systems for applications such as speaker recognition and speech recognition, using algorithms like Dynamic Time Warping (DTW), Vector Quantization (VQ), Gaussian Mixture Models	PO1, PO2, PO3, PO6, PO9, PO10

	(GMMs), Hidden Markov Models (HMMs), and Support Vector Machines (SVMs).	
7	Implement a basic speech processing system using software tools or libraries for tasks like feature extraction or speaker/speech recognition.	PO1, PO2, PO5, PO9, PO12

Topics Covered:

Unit-1: Introduction: Speech production and perception, Information sources in speech, Linguistic aspect of speech, Acoustic and articulatory phonetics, Nature of speech signal, Models for speech analysis.

Unit-2: Short-term processing: Overview of Fourier representation, Short-term Fourier transform (STFT), Filter-bank views of STFT, Time, Frequency and Time-Frequency analysis.

Unit-3: Cesptrum analysis: Basis and Development, Homomorphic signal processing, Real and Complex cepstrum, Mel-frequency cepstral coefficient (MFCC), Delta and Delta-Delta.

Unit-4: Linear Prediction (LP) analysis: Basis and Development, Levinson-Durbin's method, Normalized error, LP spectrum, LP cepstrum, LP residual.

Unit-5: Sinusoidal analysis: Basis and Development, Phase unwrapping, Sinusoidal Analysis and Synthesis of speech signal.

Unit-6: Applications of speech signal processing:

(i) Speaker recognition: Objective, Issues, Block diagram description, Classification, Development of speaker recognition system using Dynamic Time Warping (DTW), Vector Quantization (VQ), Gaussian Mixture Model (GMM) and GMM-Universal Background Model (UBM, Support Vector Machine (SVM).

(ii) Speech recognition: Objective, Issues, Block diagram description, Development of speech recognition system using DTW, VQ and Hidden Markov Model (HMM).

Text books:

1. J. R. Deller, Jr., J. H. L. Hansen and J. G. Proakis, " Discrete-Time Processing of Speech Signals",Wiley-IEEE Press, NY, USA, 1999.

2. L. R. Rabiner, B. H. Jhuang and B. Yegnanarayana, "Fundamentals of speech recognition", Pearson Education, 2009.

3. L.R. Rabiner and R.W. Schafer, "Digital Processing of Speech Signals", Pearson Education, Delhi, India, 2004

Reference books:

1. D. O'Shaughnessy, "Speech Communications: Human and Machine", Second Edition, University Press, 2005.

2. T. F. Quatieri, “Discrete time processing of speech signals”, Pearson Education, 2005.
3. J. Benesty, M. M. Sondhi and Y. Huang, “Handbook of speech processing”, Springer, 2008.