PVP SIDDHARTHA INSTITUTE OF TEHNOLOGY, KANURU, VIJAYAWADA
(AUTONOMOUS)

INFORMATION TECHNOLOGY

SPEECH PROCESSING

Course Code	19IT4602C	**Year**	III	**Semester**	II
Course Category	**Program Elective**	**Branch**	IT	**Course Type**	Theory
Credits	3	**L-T-P**	3-0-0	**Prerequisites**	AI Tools
Continuous Internal Evaluation :	30	**Semester End Evaluation:**	70	**Total Marks:**	100

Course Outcomes		**Blooms Taxonomy Level**
Upon successful completion of the course, the student will be able to		
CO1	Understanding of Fundamentals Concepts of Speech Processing	L2
CO2	Compare and Contrast on speech analysis and synthesis technologies, and their strengths and limitations	L2
CO3	Design and evaluate simple studies that utilize speech processing methods	L3
CO4	Analyze Speech Recognition models and applications	L3
CO5	Analyze Speech Synthesis applications	L3

Contribution of Course Outcomes towards achievement of Program Outcomes & Strength of correlations (3:Substantial, 2: Moderate, 1:Slight)

	PO 1	**PO 2**	**PO 3**	**PO 4**	**PO 5**	**PO 6**	**PO 7**	**PO 8**	**PO 9**	**PO1 0**	**PO1 1**	**PO1 2**	**PSO 1**	**PSO 2**
CO 1	2					2							1	
CO 2	2	3				2							1	
CO 3	2					2							1	
CO 4	2	2				2							1	
CO 5	2	2				2							1	

Syllabus

Unit No	**Contents**	**Mapped CO**
I	**Basic Concepts Speech Fundamentals**: Articulatory Phonetics – Production and Classification of Speech Sounds; Acoustic Phonetics –	CO1

	Acoustics of speech production; Review of Digital Signal Processing concepts; Short-Time Fourier Transform, Filter-Bank and LPC Methods.	
II	**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 using a Warped Frequency Scale, LPC, PLP and MFCC Coefficients, Time Alignment and Normalization – Dynamic Time Warping, Multiple Time – Alignment Paths	CO2
III	**Speech Modeling**: Hidden Markov Models: Markov Processes, HMMs – Evaluation, Optimal State Sequence – Viterbi Search, Baum-Welch Parameter Re-estimation, Implementation issues	CO3
IV	**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.	CO4
V	**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	CO5

Learning Recourses

Text Books

1. Lawrence Rabiner and Biing-Hwang Juang, "Fundamentals of Speech Recognition", Pearson Education, 2003.
2. Daniel Jurafsky and James H Martin, "Speech and Language Processing – An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition", Pearson Education, 2002.

References

1. Steven W. Smith, "The Scientist and Engineer's Guide to Digital Signal Processing", California Technical Publishing, 1997.
2. Thomas F Quatieri, "Discrete-Time Speech Signal Processing – Principles and Practice", Pearson Education, 2004.

e-Resources & other digital material

NPTEL VIDEO LECTURES