Multimodal Data Processing	**L**	**T**	**P**	Machine Learning
	3	**1**	**0**	

Course Objective:: To Understand the fundamentals of Multimodal data, text processing techniques and language models

S. NO	Course Outcomes (CO)
CO1	Identify and explain the idea of multimodal data processing along with its applications in text processing
CO2	Locate and describe various terminologies in Speech processing
CO3	Interpret and analyze different digital image and video processing approaches.
CO4	Demonstrate the need of Conventional multi-modal learning and co-learning.

S. NO	Contents	Contact Hours
UNIT 1	Introduction: Introduction to Multimodal data and applications, Multimodal Representation: two broad approaches, Joint and Coordinated. Challenges of multimodal data, Data collection & cleaning. Text Processing: Text normalization, Lemmatization, Morphology, Sub word tokenization; Text processing and statistics: TFIDF, BM-25, Zipf's law, Hipf's law; Language models and smoothing techniques; Vector space models.	**8**
UNIT 2	Speech Processing: Speech production and perception, Acoustic and articulatory phonetics; Short- term analysis: Need and windowing, Energy, Zero-crossing rate, Autocorrelation function, Fourier transform, Spectrogram; Short-term synthesis: Overlap-add method; Cepstrum analysis: Basis and development, mel-cepstrum.	**6**
UNIT 3	Digital Image and Video Processing: Point processing, Neighborhood processing, Enhancement, Edge detection, Segmentation, Feature descriptors, Restoration, Morphological operations, Image transforms, Spatial and temporal data handling	**6**
UNIT 4	Multi-modal learning and associated challenges: Applications and challenges from fusing two or more modalities such as vision, language, audio, graphs, biomedical signals, Development of shallow and deep networks for multimodal learning.	**8**
UNIT 5	Multi-modal processing and learning with applications: Image captioning, visual questioning answering system, automatic commentary generation, cognitive state estimation, recommendation system. Other Modalities: Biomedical signals, and Conventional multi-modal learning, co-learning etc.	**8**
	TOTAL	**42**