



INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

Dundigal - 500 043, Hyderabad, Telangana

COURSE CONTENT

Multi Model AI Systems

V Semester: CSE

Course Code	Category	Hours / Week			Credits	Maximum Marks		
		L	T	P		CIA	SEE	Total
ACSE42	Elective	2	0	0	2	40	60	100
		Contact Classes: 48		Tutorial Classes: Nil		Practical Classes: Nil		Total Classes: 48
Prerequisite: Information Security and Management								

I. COURSE OVERVIEW:

This course introduces the fundamentals and advanced concepts of Multimodal Artificial Intelligence Systems, focusing on the integration and processing of multiple data modalities such as text, image, audio, and video. It covers multimodal representation learning, fusion techniques, alignment methods, model architectures, evaluation strategies, and real-world applications. The course enables students to design intelligent systems that can understand and reason across multiple modalities.

II. COURSES OBJECTIVES:

The students will try to learn

- I. The fundamentals and motivation behind multimodal artificial intelligence systems.
- II. Different data modalities and their characteristics in AI systems
- III. Multimodal representation learning and feature extraction techniques.
- IV. Fusion strategies and alignment methods for combining multimodal data

III. COURSE OUTCOMES:

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

- CO1 Understand the concepts and importance of multimodal artificial intelligence systems.
- CO2 Identify and process different data modalities such as text, image, audio, and video
- CO3 Apply multimodal representation learning and feature extraction techniques.
- CO4 Analyze and implement multimodal fusion and alignment strategies
- CO5 Evaluate multimodal AI models using appropriate performance metrics.
- CO6 Develop multimodal AI solutions for real-world applications.

IV. COURSE CONTENT:

MODULE - I: INTRODUCTION TO MULTIMODAL AI (XAI) (09)

Introduction to multimodal artificial intelligence, motivation and challenges of multimodal learning, overview of data modalities, unimodal vs multimodal systems, multimodal datasets and benchmarks

MODULE – II: MULTIMODAL REPRESENTATION LEARNING (10)

Feature extraction for text, image, audio, and video, embeddings and shared representation spaces, cross-modal representation learning, self-supervised and contrastive learning methods.

MODULE – III: MULTIMODAL FUSION AND ALIGNMENT (10)

Early fusion, late fusion, hybrid fusion techniques, multimodal alignment and synchronization, attention-based fusion models, transformer-based multimodal architectures.

MODULE – IV: MULTIMODAL MODELS AND EVALUATION (10)

Multimodal deep learning architectures, training strategies, loss functions for multimodal learning, evaluation metrics, robustness and generalization in multimodal systems.

MODULE – V: APPLICATIONS OF MULTIMODAL AI SYSTEMS (9)

Multimodal AI in healthcare, multimedia analysis, human–computer interaction, autonomous systems, recommendation systems, natural language and vision applications

IV. TEXT BOOKS:

1. Pradeepta Mishra, “*Multimodal Machine Learning with Python*”, Apress, 2023.
2. Li Deng, “*Multimodal Intelligence: Representation, Learning, and Reasoning*”, Springer, 2021.

V. REFERENCE BOOKS:

1. Tadas Baltrušaitis, Chaitanya Ahuja, Louis-Philippe Morency, “*Multimodal Machine Learning: A Survey and Taxonomy*,” IEEE TPAMI, 2019.
2. Hao Tan, Mohit Bansal, “LXMERT: Learning Cross-Modality Encoder Representations from Transformers”, EMNLP, 2019

VI. ELECTRONICS RESOURCES:

1. <https://paperswithcode.com/task/multimodal-learning>
2. <https://ai.googleblog.com/search/label/Multimodal/>

VII. MATERIALS ONLINE

1. Course Outline Description
2. Lecture notes
3. PowerPoint presentation
4. Definitions and Terminology
5. Tutorial Question Bank
6. Case Studies
7. Real life Examples
8. Complex Engineering Problems
9. Tech Talk Topics
10. Concept Video Topics
11. Open-ended Exercises
12. Assignments
13. Model Question Paper – I
14. Model Question Paper – II
15. GATE Question Bank
16. Previous Question Papers and Solutions