Course Name : Multimodal AI
Course Code : TE7493
Faculty : Engineering
Course Credit : 3
Course Level : 3
Sub-Committee (Specialization) : Computer Science
Learning Objectives :

Learn the fundamental concepts associated with MMDF methods.
Apply the basic principles to derive the various multimodal techniques for different modalities such as text-text, image-text etc
Understand and choose various approaches based on the specific application for designing MMDF algorithms
Implement and validate the MMDF for specific applications.

Books Recommended :

Book	Author	Publisher
Multimodal data fusion based on mutual information	Bramon, R., Boada, I., Bardera, A., Rodriguez, J., Feixas, M., Puig, J., & Sbert, M. (2011)	IEEE Transactions on Visualization and Computer Graphics, 18(9), 1574-1587
Multimodal data fusion: an overview of methods, challenges, and prospects	Lahat, D., Adali, T., & Jutten, C. (2015)	Proceedings of the IEEE, 103(9), 1449-1477
Multimodal machine learning: A survey and taxonomy	Baltrusaitis, T., Ahuja, C., & Morency, L. P. (2018)	IEEE transactions on pattern analysis and machine intelligence, 41(2), 423-443

Course Outline :

Sr. No.	Topic	Actual Teaching Hours	Contact Hours Equivalence
1	**IntroductionWhat is Multimodal , Historical view, multimodal vs multimedia, Why multimodal, Multimodal applications: image captioning, video description, AVSR, Core technical challenges, Representation learning, translation, alignment, fusion and co-learn**	**12**	**0**
2	**Multimodal representations Joint representations, Visual semantic spaces, multimodal autoencoder, Orthogonal joint representations, Component analysis, Parallel multimodal representations, Similarity metrics, canonical correlation analysis.**	**10**	**0**
3	**Multimodal translation and mappingEncoder-decoder models, Machine translation, image captioning, Generative vs retrieval approaches, modality alignment, Latent alignment approaches, Attention models, multi instance learning, Explicit alignment, Dynamic ti**	**13**	**0**
4	**Multimodal fusion and co-learningModel free approaches, Early and late fusion, hybrid models, Kernel-based fusion, Multiple kernel learning, Multimodal graphical models, Factorial HMM, Multi-view Hidden CRF**	**10**	**0**
	Total	**45**	**0**

Pre Requisites :

Deep Learning

Evaluation :

A) Continuous Assessment (90 marks)
a)Quizzes
b)Assignments
c)Tests
B) End Semester Examination (60 Marks)

Pedagogy :

1.Online Video lectures
2.Webinars
3.Online activities

Expert :

Dr. Ketan Kotecha,Professor and Director,SIT
Dr. Shraddha Phansalkar,Associate Professor and Head (CS & IT),SIT