Course Instructor: Dr Chiranjib Sur

Course Webpage: https://chiranjibsuruf.github.io/courses/da627.html

Email: chiranjib@iitg.ac.in

Objectives: Building Multimodal GenAI course matters because the world is multimodal. Learning to fuse text, images, audio, and video lets you build AI that truly understands context and acts intelligently. This course equips you to create next-gen applications in healthcare, education, creativity, and automation that pure text models can't deliver.

Prerequisites: Machine Learning, Deep Learning, Data Mining.

Course Code: DA627

Course Name: Building Multimodal GenAI

Credits: 3-0-0-6

Syllabus: Introduction to Gen AI: LLma, GPT, Gemini, Mistral, Camel, Qwen, datasets, platforms; Text generation: training, quantification, domain adaptation, fine-tuning; Image and video generation: usage; Multimodal Gen AI: digital transformation, business opportunities; Gen AI for audio and voice: fine-tuning, deep fake; Gen AI for simulation, logistics, forecasting, disentanglement, academic integrity, education and multi-objective optimization, biology; Dynamics of reinforcement learning and data search in Generative AI;

Textbooks:

- Will be updated soon.

References:

- Will be updated soon.