MODULE-WISE COURSE OVERVIEW (30–50 Hour AI/ML Course)

1. Introduction to Artificial Intelligence

- What is AI? Definition and history
- Types of AI: Narrow, General, Superintelligent
- Real-world applications
- Overview of ML, DL, NLP, CV, GenAI
- Modern AI ecosystem: LLMs, multimodal AI

2. Machine Learning (ML)

- What is ML?
- Supervised, Unsupervised, Reinforcement learning
- Data preprocessing, feature engineering, EDA
- Regression & classification models
- Ensemble methods: Bagging, Boosting
- Clustering: K-Means, DBSCAN, Hierarchical
- Dimensionality reduction: PCA, LDA, t-SNE
- Evaluation & tuning: Grid/Random search, regularization, overfitting, bias–variance

3. Deep Learning (DL)

- Introduction to Neural Network
- Perceptron, activations, backpropagation
- Fully Connected Networks
- CNNs
- RNN, LSTM, GRU

- GNN

- TRANSFORMERS & ATTENTION

- GANs vs DIFFUSION MODELS

- LOSS FUNCTIONS, OPTIMIZERS, REGULARIZATION

---

4. NATURAL LANGUAGE PROCESSING (NLP)

- CLASSICAL NLP: TOKENIZATION, STOPWORDS, STEMMING, TF-IDF

- TEXT EMBEDDINGS (WORD2VEC, GLOVE), TOKENIZATION,

- RNN/LSTM/GRU

- ATTENTION

- CONTEXT

- TRANSFORMERS: BERT, GPT, T5

- TASKS: NER, QA, SUMMARIZATION, TRANSLATION

- PROMPT ENGINEERING

- TOOLS: NLTK, SPACY, HUGGINGFACE

- NLP PROJECTS

---

5. COMPUTER VISION (CV)

- COMPUTER VISION BASICS AND CHALLENGES

- IMAGE FORMATS, PIXELS

- CONVOLUTIONS, FILTERS, SIFT, HOG

- OBJECT DETECTION: R-CNN → YOLO → DETR

- SEGMENTATION: U-NET, MASK R-CNN

- FACE DETECTION/RECOGNITION

- VISION TRANSFORMER (VIT)

- CV PROJECTS

6. Transfer Learning & Fine-Tuning

- Why pretrained models?
- Reusing models in NLP, CV, multimodal tasks
- Fine-tuning methods: Full, Adapters, LoRA, QLoRA,...
- Prompt-tuning vs fine-tuning
- Use cases

7. Retrieval-Augmented Generation (RAG)

- Why RAG?
- RAG pipeline (embed → index → retrieve → generate)
- Vector databases (FAISS, Chroma, Milvus)
- Chunking strategies
- Retrieval evaluation
- RAG vs fine-tuning

8. Speech AI

- ASR (Whisper)
- TTS (VITS, Bark)
- Speaker identification
- Emotion recognition
- Voice assistants
- Speech + LLM multimodal agents

9. Graph Neural Networks (GNNs)

Overview

- What are Graph Neural Networks and why do we need them?
- Representing data as graphs (nodes, edges, adjacency matrices)
- Graph learning vs traditional ML/DL
- Applications: social networks, recommendation systems, fraud detection, molecular predictions, knowledge graphs

Core Concepts

- Graph convolution
- Message passing
- Neighborhood aggregation
- Graph embeddings

Popular GNN Architectures

- GCN (Graph Convolutional Network)
- GAT (Graph Attention Network)
- GraphSAGE
- Relational GCN (R-GCN)
- Heterogeneous graph models

Advanced GNN Topics

- Link prediction
- Node classification
- Graph classification
- Knowledge graph embeddings
- Combining LLMs + GNNs (emerging field)

Tools

- PyTorch Geometric (PyG)
- DGL (Deep Graph Library)

GNN Projects

- Social network node classification

- Molecular property prediction

- Fraud detection using transaction graphs

---

10. Advanced Topics

- DETR

- Multimodality

- decoding strategies

- MoE

- Large Language Models: Architecture, training, applications (ChatGPT, LLaMA, Mistral)

---

11. Vision–Language Models (VLMs)

- What are VLMs?

- CLIP

- GroundingDINO (language-guided object detection)

- SAM (foundation segmentation model)

- Grounded-SAM workflow

- Applications: multimodal search, robotics, VQA

---

12. AI Explainability

- Explainability: Grad-CAM, attention viz

13. End-to-End Projects

- ML pipeline

- CNN classifier

- NLP (NER, sentiment)

- Transformer-based chatbot

- RAG-based system

- GroundingDINO + SAM detector

- Whisper ASR tool

- LLM LoRA/QLoRA fine-tuning