



ADVANCED GENERATIVE AI

LARGE LANGUAGE MODELS
EVAL AUGMENTED GENER
E-TUNING · MULTIMODAL

LEVEL	MODULES	MODE	DELIVERY
Advanced	18	Classroom · Online · Hybrid	Lecture · Lab · Projects

01 COURSE OVERVIEW

This programme takes learners from the foundations of artificial intelligence to production-grade generative AI systems. It covers the mathematics and Python engineering that underpin modern models, the architecture of large language models, and the applied disciplines that surround them: prompt engineering, retrieval-augmented generation, agent design, fine-tuning, multimodal generation, evaluation, security and MLOps.

Every module pairs concepts with a lab on the tools used in industry — Hugging Face, LangChain, LlamaIndex, vector databases, FastAPI and Docker — and the programme closes with an end-to-end capstone build.

WHO IT IS FOR

- Software developers moving into AI engineering
- Data science and analytics professionals
- Final-year engineering and MCA students
- Technical leads evaluating AI for their products

PREREQUISITES

Basic programming knowledge. Python fundamentals recommended; the programme revises them in Module 3.

02 LEARNING OUTCOMES

- | | |
|---|---|
| 01 Explain how transformer-based language models are built, scaled and served. | 02 Write, chain and optimise prompts for reliable production behaviour. |
| 03 Build retrieval pipelines over private documents using vector databases. | 04 Design autonomous and multi-agent systems with tool and function calling. |
| 05 Fine-tune open models with LoRA, QLoRA and PEFT on prepared datasets. | 06 Generate and integrate image, audio and speech output alongside text. |
| 07 Deploy AI applications as APIs and containers and monitor them in use. | 08 Evaluate output with BLEU, ROUGE, perplexity and LLM-as-a-judge methods. |
| 09 Identify and mitigate prompt injection, bias, privacy and governance risks. | |
-

03 MODULE STRUCTURE

18 MODULES

01 FOUNDATIONS OF ARTIFICIAL INTELLIGENCE

Evolution of AI

AI Applications across Industries

History of Large Language Models (LLMs)

Machine Learning vs Deep Learning vs Generative AI

AI Ecosystem

AI Ethics and Responsible AI

LAB

SETTING UP AI DEVELOPMENT ENVIRONMENT

INTRODUCTION TO GOOGLE COLLAB, JUPYTER NOTEBOOK

USING PYTHON FOR AI

02 MATHEMATICS FOR GENERATIVE AI

Linear Algebra

Calculus Basics

Gradient Descent

Probability & Statistics

Optimization

Loss Functions

LAB

NUMPY

MATRIX OPERATIONS

DATA VISUALIZATION USING MATPLOTLIB

03 PYTHON FOR AI DEVELOPMENT

Python Fundamentals
Functions
File Handling
Virtual Environments

Data Structures
OOP
Exception Handling

LAB

- PANDAS
- NUMPY
- MATPLOTLIB
- SEABORN
- SCIKIT-LEARN BASICS

04 NATURAL LANGUAGE PROCESSING (NLP)

Text Processing
Stemming
Word Embeddings
Word2Vec
BERT Introduction

Tokenization
Lemmatization
TF-IDF
FastText

LAB

- NLTK
- SPACY
- HUGGING FACE TOKENIZERS

05 DEEP LEARNING FUNDAMENTALS

Artificial Neural Networks
RNN
GRU
Attention Mechanism

CNN
LSTM
Transformers

LAB

- TENSORFLOW
- KERAS
- PYTORCH BASICS

06 LARGE LANGUAGE MODELS (LLMS)

GPT Architecture
Encoder vs Decoder
Attention Layers
Context Windows

Transformer Architecture
Positional Encoding
Token Prediction
Model Scaling

POPULAR MODELS

- GPT-4/5
- CLAUDE
- GEMINI
- LLAMA
- DEEPSEEK
- MISTRAL
- QWEN

LAB

- HUGGING FACE TRANSFORMERS
- MODEL INFERENCE

07 PROMPT ENGINEERING (ADVANCED)

Zero-shot Prompting

Few-shot Prompting

Tree-of-Thought

System Prompt Design

Prompt Optimization

One-shot Prompting

Chain-of-Thought Prompting

Role Prompting

Prompt Chaining

LAB

CHATGPT

CLAUDE

GEMINI

PERPLEXITY

MINI PROJECTS

AI TUTOR

AI HR ASSISTANT

AI CUSTOMER SUPPORT BOT

08 RETRIEVAL-AUGMENTED GENERATION (RAG)

Why RAG?

Chunking

Document Loading

Embeddings

VECTOR DATABASES

SEMANTIC SEARCH

RETRIEVAL PIPELINES

CONTEXT INJECTION

VECTOR DATABASES

FAISS

CHROMADB

PINECONE

WEAVIATE

LAB

LANGCHAIN

LLAMAINDEX

CHROMADB

PROJECT

PDF CHATBOT

COMPANY KNOWLEDGE ASSISTANT

09 AI AGENTS

Agent Architecture

Planning

Tool Calling

Multi-Agent Systems

Autonomous Agents

Memory

Function Calling

FRAMEWORKS

LANGGRAPH

CREWAI

AUTOGEN

OPENAI AGENTS SDK

LAB

BUILD AI RESEARCH AGENT

TRAVEL PLANNER AGENT

EMAIL AUTOMATION AGENT

10 FINE-TUNING LLMS

Supervised Fine-Tuning (SFT)

LoRA

PEFT

DPO

Instruction Tuning

QLoRA

RLHF

LAB

FINE-TUNING WITH HUGGING FACE

DATASET PREPARATION

11 MULTIMODAL AI

Text Generation

Video Generation

Speech Models

Image Generation

Audio Generation

Vision Models

MODELS

DALL-E

STABLE DIFFUSION

FLUX

MIDJOURNEY

WHISPER

ELEVENLABS

LAB

IMAGE GENERATION

SPEECH-TO-TEXT

TEXT-TO-SPEECH

12 AI APPLICATION DEVELOPMENT

REST APIs

Flask

Gradio

Deployment Basics

FastAPI

Streamlit

Docker

PROJECT

AI WEB APPLICATION

13 AI AUTOMATION

Workflow Automation

Email Automation

OCR Integration

AI Assistants

Document Automation

Calendar Automation

TOOLS

ZAPIER

N8N

MAKE

POWER AUTOMATE

14 AI SECURITY & RESPONSIBLE AI

AI Risks

Jailbreaking

Model Bias

AI Governance

Prompt Injection

Hallucinations

Data Privacy

Secure AI Deployment

15 MODEL EVALUATION

BLEU

Perplexity

LLM-as-a-Judge

Cost Optimization

ROUGE

Human Evaluation

Benchmarking

16 MLOPS FOR GENERATIVE AI

Model Versioning

CI/CD

Scaling

Experiment Tracking

Monitoring

GPU Deployment

TOOLS

MLFLOW

WEIGHTS & BIASES

DOCKER

KUBERNETES

GITHUB ACTIONS

17 CLOUD AI SERVICES

OpenAI API

Google Vertex AI

Hugging Face Inference API

Azure OpenAI

AWS Bedrock

LAB

API INTEGRATION

SECURE KEY MANAGEMENT

COST MONITORING

18 CAPSTONE PROJECTS

Students build end-to-end Generative AI solutions, such as:

AI Legal Assistant

AI Resume Analyzer

AI Code Generator

AI Customer Support Bot

AI Research Assistant

AI Sales Assistant

Enterprise RAG Chatbot

AI Medical Assistant

AI Interview Coach

AI Content Creator

AI Document Summarizer

AI Voice Assistant

04 TOOLS & TECHNOLOGIES COVERED

RECOMMENDED TOOLS

PYTHON

VS CODE

GOOGLE COLAB

JUPYTER NOTEBOOK

GIT & GITHUB

HUGGING FACE

LANGCHAIN

LLAMA INDEX

CHROMADB

FAISS

STREAMLIT

GRADIO

FASTAPI

DOCKER

OPENAI API

05 ASSESSMENT & CERTIFICATION

Module assignments and labs	25%
Mini projects	20%
Internal assessments	15%
Capstone project and review	40%

CERTIFICATION

Learners who complete all modules, submit the capstone project and clear the final review receive a course completion certificate from Kaizen Infinities Private Limited. Project work is documented for the learner's portfolio and GitHub profile, and interview preparation is included in the final week.

06 CAREER OUTCOMES

ROLES THIS PROGRAMME PREPARES FOR

GENERATIVE AI ENGINEER

LLM APPLICATION DEVELOPER

AI AGENT DEVELOPER

PROMPT ENGINEER

NLP ENGINEER

AI PRODUCT ENGINEER

MLOPS ENGINEER

AI AUTOMATION CONSULTANT

CALL US

+91 73051 33996

EMAIL US

apply@kaizeninfinities.com

VISIT OUR WEBSITE

www.kaizeninfinities.com