Superior University Gold Campus

Department Of Software Engineering

Submitted By:

Ahmar Tariq

Roll No:

SU92-BSAIM-F23-111

Subject:

Programming For Artificial Intelligence (Lab)

Submitted To:

Sir Rasikh Ali

Lang Chain:

- A framework to build applications using LLMs.
- **Purpose**: Helps developers chain together LLMs with tools like memory, search, databases, APIs, etc.
- **Use case:** Build chatbots, agents, or document question-answering systems that can interact with external data.

RAG (Retrieval-Augmented Generation):

- An architecture that combines information retrieval with LLM generation.
- **Purpose:** Helps LLMs give better answers by pulling relevant facts from external data before generating text.
- **How it works:**
    1. **Retrieve** relevant documents (e.g. from a Vector DB).
    2. **Augment** the LLM prompt with those docs.
    3. **Generate** a response using the LLM.
- **Use case:** Search-based chat, customer support, legal or medical assistants.

LLMs (Large Language Models):

- AI models trained on vast amounts of text to understand and generate human-like language.
- **Examples:** GPT-4, Claude, LLaMA, PaLM.
- **Use case:** Text generation, summarization, code writing, translation, chatbots, etc.

FAISS (Facebook AI Similarity Search):

- A **library** for fast similarity search over dense vectors (embeddings).
- **Purpose:** Helps find the closest match (e.g., similar documents) from a big set of vectorized data.
- **Use case:** Powering the retrieval step in RAG, search engines, recommendation systems.

Vector:

- A numerical representation of data (e.g., a sentence) in multi-dimensional space.
- **Why it matters:** Similar things (e.g. similar questions or documents) are close together in vector space.
- **Use case:** Comparing semantic similarity, powering search and recommendation systems.

Vector DB (Vector Database):

- A database optimized for storing and querying vectors.
- **Examples:** Pinecone, Weaviate, Qdrant, Milvus.
- **Use case:** Fast retrieval of semantically similar items (e.g., docs, images) using embeddings.

Generative AI:

- A broad field of AI that generates new content — text, images, music, code, etc.

- **Includes:** LLMs (text), GANs (images), diffusion models (images), etc.
- **Use case:** Chatbots, art generation, video generation, code completion, music composition.

GANs (Generative Adversarial Networks):

- A type of generative model using two neural networks: Generator and Discriminator.
- **Purpose:** Generator creates fake data, Discriminator tries to detect fakes — they compete until the generator becomes good.
- **Use case:** Creating realistic images, deepfakes, image upscaling, art generation.