**Lab 11 task**

**1.**

**LangChain** is a Python framework designed to simplify the process of building applications powered by large language models (LLMs). It enables developers to link various components such as memory, tools, retrievers, and chains of prompts to create intelligent systems like chatbots, Q&A systems, and document analysis tools. LangChain is particularly useful when an LLM needs to interact with external tools or structured data, making it easier to build end-to-end generative AI applications.

**2.**

**RAG (Retrieval-Augmented Generation)** is a technique that enhances the performance of language models by combining them with external knowledge retrieval. Instead of generating answers solely based on pre-trained knowledge, RAG retrieves relevant documents from a database or corpus and provides them as context to the LLM. This improves the accuracy and reliability of responses, especially when answering domain-specific or recent questions, and reduces the risk of generating hallucinated content.

**3.**

**LLMs (Large Language Models)** are advanced AI models trained on massive datasets of human language to understand, generate, and manipulate text. Examples include OpenAI's GPT-3 and GPT-4, Meta's LLaMA, and Google's BERT. These models are capable of performing a variety of natural language tasks such as translation, summarization, question answering, and content creation. Their versatility has made them central to modern AI applications.

**4.**

**FAISS (Facebook AI Similarity Search)** is a library developed by Facebook AI for efficient similarity search and clustering of dense vectors. It is widely used in applications where fast and scalable retrieval of similar items (like documents or images) is needed. FAISS excels at handling high-dimensional data and is a core component in vector search engines or vector databases used in RAG systems.

**5.**

A **Vector** in the context of AI is a numerical representation of data, often in a high-dimensional space, that captures the semantic meaning of inputs like text, images, or audio. These vectors are generated using embedding techniques and allow the system to compare similarity between different inputs mathematically. Vectors are fundamental in tasks like recommendation, clustering, and information retrieval.

**6.**

**A VectorDB (Vector Database)** is a specialized type of database optimized to store and retrieve vector embeddings based on similarity. Instead of traditional queries, it supports operations like nearest-neighbor search, which is essential for applications like document search, image retrieval, and chatbot memory. Popular VectorDBs include FAISS, Pinecone, Chroma, and Weaviate, which are often used in combination with LLMs in RAG setups.

**7.**

**Generative AI** refers to a category of artificial intelligence that can create new content such as text, images, music, and even code. It uses models like GPT, DALL·E, and MusicLM, which learn patterns from large datasets and generate outputs that resemble human-created content. Generative AI has revolutionized creative fields and is also used in automation, design, writing assistants, and entertainment.

**8.**

**GANs (Generative Adversarial Networks)** are a class of deep learning models consisting of two neural networks—a Generator and a Discriminator—that compete with each other. The Generator creates fake data while the Discriminator tries to distinguish it from real data. This adversarial process pushes the Generator to produce increasingly realistic outputs. GANs are particularly known for their ability to generate lifelike images, deepfakes, and synthetic media.