

AI/ML Trainee -Assignment 2

Name: Mahendra kolipaka

Question 1: Self Assessment

Machine Learning (ML): B - I understand core ML concepts and can implement basic models with guidance.

Artificial Intelligence (AI): B - I have foundational knowledge of AI and have worked on small projects.

Deep Learning (DL): C - I have basic theoretical understanding but limited practical experience.

Large Language Models (LLMs): C - I understand the high-level working of LLMs and prompt engineering.

Question 2: LLM-based Chatbot Architecture

An LLM-based chatbot consists of several key components:

1. User Interface - Allows users to interact with the chatbot .
2. Backend Server - Handles requests and manages chatbot logic.
3. Large Language Model - Generates responses based on user input.
4. Prompt Engineering - Structures input to improve response quality.
5. Context Management - Stores conversation history for continuity.
6. External APIs and Tools - Used for fetching data or performing actions.
7. Safety and Moderation Layer - Ensures safe and reliable responses.

Question 3: Vector Databases

Vector databases are used to store and search vector embedding generated by machine learning models. These embeddings represent text or data numerically and allow semantic search.

Vector databases are commonly used in recommendation systems, semantic search, and Retrieval - Augmented Generation (RAG).

If I were to choose a vector database, I would select FAISS because it is open- source, efficient, and suitable for Large -scale similarity search .For cloud -based systems, Pinecone is a good option due to scalability.