Assignment 2: LLM, AI, ML Self-Assessment and System Understanding

1. Self-Assessment on LLM, Deep Learning, AI, and Machine Learning:

This section presents an honest and structured self-assessment of my current skill level across key domains related to Artificial Intelligence and Machine Learning. The assessment uses the following scale:

A – Can code independently

B – Can code under supervision

C – Little or no understanding

Area	Rating	Explanation
Large Language Models (LLM)	B	I understand core LLM concepts such as transformer architecture, embeddings, attention mechanisms, and prompt engineering. I can build LLM-based workflows with guidance and am actively improving my independent implementation skills.
Deep Learning	A	I have strong hands-on experience designing and implementing deep learning models using TensorFlow. I have independently built CNN-based and NLP-based models, including training, tuning, and evaluation.
Artificial Intelligence (AI)	B	I have a solid understanding of fundamental AI concepts and can develop AI-driven systems under supervision. I am progressing toward architecting complete end-to-end AI solutions independently.
Machine Learning (ML)	A	I have extensive experience working on real-world machine learning problems involving data preprocessing, feature engineering, model training, evaluation, and deployment. I can independently deliver ML solutions.

2. High-Level Architecture of an LLM-Based Chatbot

An LLM-based chatbot is designed to understand user input in natural language and generate meaningful, context-aware responses. The system consists of several interconnected components working together.

1**. User Interface:** The front-end layer where users interact with the chatbot through text input, such as a web or mobile application.

2**. Backend Service / API Layer:** Acts as the controller that processes user input, manages workflows, and communicates with the LLM and databases.

3. **Large Language Model (LLM):** Responsible for understanding intent, context, and generating human-like responses.

4. **Knowledge Source:** Domain-specific documents such as FAQs, manuals, or policy documents that provide factual grounding.

5. **Vector Database:** Stores embeddings of documents to enable semantic search and retrieval of relevant information.

6. **Response Generation Layer:** Combines the retrieved context with the user query to produce an accurate and coherent response.

7. **Example:**

In a banking chatbot, when a user asks 'How do I block my debit card?', the system retrieves relevant policy documents from the vector database and uses the LLM to generate a clear, user-friendly response.

3. Vector Databases:

A vector database is a specialized database designed to store and search high-dimensional vectors (embeddings). These embeddings capture the semantic meaning of unstructured data such as text or images. Unlike traditional databases that rely on keyword matching, vector databases enable semantic search, allowing systems to retrieve results based on meaning rather than exact word matches.

Hypothetical Use Case and Vector Database Selection Use Case:

A customer support chatbot for a software product that answers user queries using product documentation. Chosen Vector Database: FAISS. Reason: FAISS provides efficient similarity search, integrates well with Python-based ML workflows, and is suitable for medium-scale deployments without requiring heavy infrastructure

4. Summary of Learning Outcomes:

The assignment reinforced understanding of:

1. **Self-Assessment Accuracy:** Ratings reflect genuine expertise validated through internships and projects.
2. **Architecture Design:** Understanding how to compose multiple components (UI, backend, LLM, knowledge sources) into coherent systems.
3. **Technology Selection:** Making informed choices based on requirements, scale, cost, and deployment flexibility.
4. **Vector Databases:** Recognizing their critical role in modern AI systems and ability to evaluate solutions systematically.

Key Insights:

- **On LLM systems:** They are not standalone solutions; they require careful orchestration with knowledge sources, retrieval systems, and monitoring to ensure accuracy and reliability in production.
- **On vector databases:** They are foundational infrastructure for AI, not optional. Choosing the right solution requires balancing performance, cost, flexibility, and feature richness based on specific use cases.
- **On continuous learning:** The AI/ML landscape evolves rapidly. B-rated skills in LLM and AI are progressing toward A as practical projects are built and production systems are understood more deeply.