AI for Technical Manager or Solution Architect

Duration - 6 Days / 48 Hours

Program Description

This program equips technical managers and solution architects with the knowledge and skills to make informed decisions about AI model selection, evaluation, and deployment. It covers critical concepts such as Retrieval-Augmented Generation (RAG) vs. Fine-Tuning/PEFT, processing raw data with LLMs, and designing agent- based architectures. Participants will also explore vectorization techniques, top vector database solutions, and best practices for building scalable AI-powered APIs and LangChain pipelines. The course balances technical depth with strategic considerations to ensure AI-driven solutions are efficient, effective, and aligned with business objectives.

Learning Goals

- Assess AI models based on performance, use case fit, and deployment strategies.
- Understand when to use Retrieval-Augmented Generation vs. Fine-Tuning techniques.
- Utilize Generative AI for data transformation, preprocessing, and enrichment.
- Implement AI-driven agents that autonomously interact with data and applications.
- Leverage vector embeddings and top vector database solutions for AI applications.

Course Topics

- Model Selection / Eval/Deployment
- RAG vs. Fine-Tuning/PEFT
- Processing raw data with GenAI/LLMs
- Agent-based arch
- Vectorization and Top Vector DB solutions
- Building API / LangChain Pipelines using LLMs