Cloud Native AI

Duration: 02 days

Courseware: Official

Lab: Without LOD

Table of Content

Module 1: Azure Cloud AI Ecosystem (4 hours)

1. **Introduction to Azure AI and Machine Learning**
    - Overview of Azure AI/ML services (Azure Machine Learning, Cognitive Services, Bot Service, etc.)
    - Use cases and applications of generative AI on Azure
2. **Azure OpenAI Service**
    - Utilizing Azure OpenAI for building generative models and applications
    - Key features and integration options with Azure AI services
3. **Integration and Deployment on Azure**
    - Building CI/CD pipelines for AI models with Azure DevOps
    - Strategies for monitoring, scaling, and maintaining models in production
4. **AI Governance and Security on Azure**
    - Responsible AI practices with Azure tools
    - Implementing data privacy, governance, and compliance strategies

Module 2: AWS Cloud AI Ecosystem (4 hours)

1. **Introduction to AWS AI/ML Services**
    - Overview of AWS AI/ML offerings (Amazon SageMaker, Comprehend, Rekognition, Polly, Translate, etc.)
    - Real-world use cases of generative AI with AWS
2. **Building AI Models with Amazon SageMaker**
    - End-to-end model development with SageMaker
    - Integration with other AWS services and best practices for model deployment
3. **Serverless AI and Machine Learning with AWS**

- Utilizing AWS Lambda and Step Functions for AI/ML workflows
- Best practices and case studies for scalable AI solutions

4. **AI Security and Governance in AWS**
    - Managing data privacy, security, and cost optimization for AI workloads
    - AWS Well-Architected Framework for AI/ML

Module 3: Google Cloud AI Ecosystem (4 hours)

1. **Introduction to Google Cloud AI Tools**
    - Overview of Google Cloud AI tools (Vertex AI, AutoML, AI Hub, etc.)
    - Generative AI applications and use cases on Google Cloud
2. **Building and Deploying Models with Vertex AI**
    - Overview of training, tuning, and deploying a generative AI model with Vertex AI
    - Integration with BigQuery, Dataflow, and other Google services for AI-driven analytics
3. **Advanced AI Capabilities in Google Cloud**
    - Leveraging AI Explanations and What-If Tool for model interpretability
    - ML Ops in Google Cloud for monitoring and managing AI models in production
4. **AI Governance, Security, and Compliance in Google Cloud**
    - Ensuring responsible AI practices with Google Cloud tools
    - Strategies for AI governance, privacy, and compliance

Module 4: NVIDIA AI Ecosystem (4 hours)

1. **Introduction to NVIDIA AI Tools and Frameworks**
    - Overview of NVIDIA AI tools (NVIDIA NGC, TensorRT, RAPIDS, etc.)
    - AI model acceleration using NVIDIA GPUs and frameworks
2. **Deploying AI Models on NVIDIA Platforms**
    - Building and optimizing AI models with NVIDIA's TensorRT
    - Applications of NVIDIA tools in cloud environments for AI acceleration
3. **NVIDIA AI at the Edge**
    - Overview of NVIDIA Jetson platform for edge AI

- Deploying AI models for edge applications with NVIDIA Jetson

4. **AI Performance Optimization and Best Practices**
    - Optimizing AI workloads for performance and scalability
    - Best practices for deploying AI models on cloud and edge devices with NVIDIA