Deploying Cloud-Based Artificial Intelligence Solutions Training Course

30 Nov - 04 Dec 2026

Kuala Lumpur

5200 € (Per Person)

Ref: #AI2643_173894

Course Introduction / Overview:

This course provides a comprehensive exploration of designing, deploying, and managing artificial intelligence solutions within cloud computing environments. As organizations increasingly migrate their AI workloads to the cloud to leverage scalability, flexibility, and powerful managed services, the demand for professionals who can bridge the gap between data science and cloud engineering has skyrocketed. This program, offered by BIG BEN Training Center, is meticulously structured to transform participants from understanding theoretical AI concepts to mastering practical cloud-native AI implementation. We will delve into the architectures and services of major cloud providers, focusing on building robust, scalable, and cost-effective AI systems. The curriculum is influenced by foundational principles of distributed systems, as detailed by authors like Martin Kleppmann in his seminal work "Designing Data-Intensive Applications," which provides a strong basis for understanding the infrastructure that underpins modern cloud AI. Participants will learn to navigate the entire MLOps lifecycle, from data ingestion and model training to automated deployment, monitoring, and governance, ensuring they can deliver real-world business value through intelligent cloud solutions. This training course is your definitive guide to becoming a proficient architect of cloud-based AI.

Target Audience / This training course is suitable for:

- AI Engineers and Machine Learning Engineers.
- Data Scientists and Data Analysts.
- Cloud Solutions Architects and Cloud Engineers.
- DevOps and MLOps Professionals.
- Software Developers and Engineers interested in AI.
- IT Managers and Project Managers overseeing AI initiatives.
- Technical Leaders and Technology Strategists.

Target Sectors and Industries:

- Information Technology and Software Services.
- Financial Services, Banking, and Insurance (FinTech).
- Healthcare and Life Sciences.
- Retail and E-commerce.
- Manufacturing and Industrial Automation.
- Telecommunications and Media.
- Government and Public Sector Agencies.
- Energy and Utilities.

Target Organizations Departments:

- Information Technology (IT) and Cloud Infrastructure.
- Research and Development (R&D).
- Data Science and Analytics.
- Software Engineering and Application Development.
- Product Management and Innovation.
- Operations and Business Intelligence.
- Digital Transformation and Strategy.

Course Offerings:

By the end of this course, the participants will have able to:

- Design scalable and resilient architectures for AI and ML workloads on the cloud.
- Evaluate and select the appropriate cloud AI services for specific business problems.
- Implement automated MLOps pipelines for continuous integration, delivery, and training.
- Deploy machine learning models using various strategies like containerization and serverless functions.
- Manage and monitor the performance, cost, and security of deployed AI solutions.
- Integrate AI models with existing applications and business processes via APIs.
- Apply best practices for data governance, security, and compliance in cloud AI environments.
- Optimize AI workloads for both performance and cost-efficiency in the cloud.

Course Methodology:

The training methodology at BIG BEN Training Center is designed to be immersive, practical, and highly interactive, ensuring participants gain tangible skills. This course moves beyond theoretical lectures to focus on hands-on application and real-world problem-solving. A significant portion of the training is dedicated to practical labs and guided projects where participants will work directly within major cloud platform environments. We employ a blended learning approach that combines expert-led instruction with collaborative group work, encouraging peer-to-peer learning and knowledge sharing. Case studies from various industries will be analyzed to demonstrate successful cloud AI deployments and common pitfalls to avoid. Interactive sessions, Q&A panels, and live demonstrations will be used to clarify complex topics like MLOps automation and serverless AI architecture. Participants will receive continuous feedback from the instructor throughout the course, helping them refine their understanding and technical skills. This hands-on, project-based approach ensures that attendees leave the course not just with knowledge, but with the confidence and competence to deploy and manage sophisticated AI solutions in their own organizations.

Course Agenda (Course Units):

Unit One: Foundations of Cloud-Based Artificial Intelligence

- Introduction to AI and Machine Learning in the Cloud.
- Overview of Major Cloud Platforms (AWS, Azure, GCP) and their AI Services.
- Understanding the Cloud AI Stack: IaaS, PaaS, and SaaS Models.
- Core Concepts: Compute, Storage, Networking, and Databases for AI.
- Architecting for Scalability, Availability, and Fault Tolerance.
- Data Ingestion and Preparation Strategies in the Cloud.
- Introduction to MLOps: Bridging Data Science and DevOps.

Unit Two: Architecting and Designing Cloud AI Solutions

- Designing Data Pipelines for Machine Learning.
- Selecting Appropriate Cloud Services for AI Workloads.
- Architectural Patterns for Training and Inference.
- Microservices Architecture for AI Applications.
- Security by Design: Identity, Access Management, and Data Encryption.
- Compliance and Data Governance Frameworks in the Cloud.
- Cost Modeling and Optimization Strategies for AI Projects.

Unit Three: Developing and Training Models in the Cloud

- Utilizing Managed Machine Learning Platforms (e.g., Amazon Sage Maker, Azure ML).
- Leveraging AutoML for Accelerated Model Development.
- Distributed Training Strategies for Large-Scale Models.
- Working with Pre-trained Models and APIs (Vision, NLP, Speech).
- Environment Management using Containers (Docker).
- Version Control for Data, Code, and Models (Git, DVC).
- Building Custom Training Environments and Algorithms.

Unit Four: Deploying and Operationalizing AI Models

- Model Deployment Strategies: Batch, Real-time, and Streaming Inference.
- Container Orchestration with Kubernetes for AI Workloads.
- Serverless Deployment using Cloud Functions (e.g., AWS Lambda).
- Implementing CI/CD Pipelines for Automated Model Deployment.
- API Gateway and Management for AI Model Endpoints.
- A/B Testing and Canary Deployments for Model Updates.
- Monitoring Model Performance, Drift, and Data Quality.

Unit Five: Advanced Topics in Cloud AI Management

- Advanced Security for AI/ML Workloads.
- AI Governance and Ethical Considerations in the Cloud.
- Cost Management and FinOps for AI.
- Edge Computing and Hybrid Cloud AI Deployments.
- Federated Learning Concepts and Use Cases.
- Future Trends in Cloud-Native AI and Generative AI.
- Capstone Project: Deploying an End-to-End AI Solution on the Cloud.

FAQ:

Qualifications required for registering to this course?

There are no requirements.

How long is each daily session, and what is the total number of training hours for the course?

This training course spans five days, with daily sessions ranging between 4 to 5 hours, including breaks and interactive activities, bringing the total duration to 20 - 25 training hours.

Something to think about:

As AI models become increasingly integrated into critical cloud infrastructure, how do we balance the need for rapid deployment with the ethical imperatives of algorithmic transparency and fairness?

What unique qualities does this course offer compared to other courses?

This training course distinguishes itself by offering a holistic, end-to-end perspective on the AI lifecycle, focusing specifically on the operational and deployment challenges that are often overlooked. While many courses concentrate solely on model building and algorithms, our curriculum is uniquely structured to bridge the critical gap between data science and cloud operations, a discipline now known as MLOps. We emphasize practical, real-world deployment strategies, from containerization with Docker and Kubernetes to serverless architectures, providing participants with a versatile and in-demand skill set. The course moves beyond the "what" to explain the "why" and "how" of architecting robust, scalable, and secure AI solutions. Rather than being tied to a single tool, it teaches architectural patterns and best practices applicable across major cloud providers, ensuring the knowledge gained is both transferable and future-proof. By integrating modules on cost optimization (FinOps), governance, and security from the outset, we equip participants not just to build models, but to deliver sustainable, enterprise-grade AI solutions that provide tangible business value. This pragmatic, operations-focused approach makes it an invaluable experience for professionals aiming to lead AI implementation in their organizations.