Hyperscaler - GCP for Data Engineers and Data Scientist

Duration - 6 Days / 48 Hours

Program Description

This 2-track, 48-hour immersive program equips Data Engineers and Data Scientists with critical knowledge and skills to build, manage, and deploy scalable data and machine learning solutions using Google Cloud Platform (GCP).

The program emphasizes hands-on labs and real-world applications with services like BigQuery, Cloud SQL, Dataflow, Dataproc, Vertex AI, and AutoML

Program Structure
GCP for Data Engineers – 24 Hours

Overview

Master GCP storage, relational and NoSQL databases, big data processing, orchestration, and pipeline development.

Learning Goals

- Understand GCP cloud storage systems (GCS, Big Table)
- Set up and manage Cloud SQL for structured data
- Build scalable pipelines using Dataflow and Dataproc
- Orchestrate workflows using Data Pipelines and Cloud Functions
- Analyze big data efficiently using BigQuery
- Integrate real-time streaming with Pub/Sub
- Create insightful dashboards with Looker and Data Studio

Course Topics

- Google Cloud Storage and Cloud SQL
- Cloud Bigtable and BigQuery
- Dataproc for Big Data Processing
- Data Pipelines and Cloud Functions
- Google Dataflow for Stream and Batch Processing
- Pub/Sub Messaging and Real-time Integration
- Looker and Data Studio for Reporting

Program Structure
GCP for Data Scientists & ML Engineers – 24 Hours

Overview

Learn how to build, train, and deploy ML models using Google Cloud services like Vertex AI, AutoML, and MLOps tools.

Learning Goals

- Explore AutoML and Vertex AI for rapid model development
- Train, evaluate, and deploy ML models using Vertex AI
- Utilize BigQuery ML for SQL-based model building
- Manage data transformation and preparation pipelines
- Monitor and version ML models effectively
- Develop end-to-end ML workflows integrated with GCP DevOps and MLOps tools

Course Topics

- Introduction to Google Cloud Machine Learning
- AutoML for No-code/Low-code Model Development
- Vertex AI for End-to-End ML Lifecycle
- Model Deployment and MLOps on GCP
- BigQuery ML for Predictive Analytics
- Monitoring ML Models and Performance
- DevOps and CI/CD Pipelines for ML