



Gemini in BigQuery for Data Practitioners

Course description

In this course, data practitioners will gain comprehensive knowledge and hands-on experience in leveraging Google's Gemini large language models directly within BigQuery. You will learn to harness BigQuery ML's capabilities to integrate generative AI functions, enabling advanced data analysis, transformation, enrichment, and insight generation. This course covers practical applications, prompt engineering best practices, responsible AI considerations, and pipeline integration strategies to empower data professionals to build innovative, AI-driven solutions within their existing BigQuery environments.

- Course level: Advanced Practitioner
- Duration: 2 days

Activities

This course includes detailed presentations, extensive hands-on labs within BigQuery, live demonstrations of advanced techniques, and group discussions on real-world use cases.

Course objectives

In this course, you will learn to:

- Explain the architecture and capabilities of Gemini models integrated with BigQuery.
- Utilize BigQuery ML to create and manage remote models for Gemini.
- Apply `GENERATE_TEXT`, `GENERATE_EMBEDDING`, and other Gemini AI functions directly within BigQuery SQL queries.
- Master prompt engineering techniques for optimal results from Gemini models in BigQuery.
- Design and implement advanced data enrichment, summarization, and semantic search solutions.
- Integrate Gemini-powered BigQuery processes into automated data pipelines.
- Understand and apply responsible AI principles for generative AI use in BigQuery.
- Optimize query performance and manage costs associated with Gemini AI functions.

Intended audience

This course is intended for data professionals who are proficient with BigQuery and seek to integrate generative AI capabilities into their data workflows:

- Data Analysts
- Data Scientists
- Machine Learning Engineers
- BigQuery Developers
- Data Architects
- Business Intelligence Developers

Prerequisites

We recommend that attendees of this course have:

- Strong proficiency in BigQuery SQL and data warehousing concepts.
- Basic understanding of machine learning concepts (e.g., embeddings, classification).
- Familiarity with Google Cloud Platform (GCP) console and services is beneficial.

Course outline

Module 1: Introduction to Gemini and BigQuery ML

- Overview of Google's Gemini family of models
- Introduction to BigQuery ML for AI/ML integration
- Why use Gemini directly in BigQuery? Use cases and benefits.
- Available Gemini models and their specific applications in BigQuery.
- Demo: Setting up BigQuery ML remote models for Gemini.

Module 2: Core Gemini AI Functions in BigQuery SQL

- `GENERATE_TEXT`: Free-form text generation, summarization, classification.
- `GENERATE_EMBEDDING`: Creating vector representations for semantic search and similarity.
- `GENERATE_QA`: (If available/applicable) Question answering from structured data.
- Applying AI functions on BigQuery tables and views.
- Hands-on Lab: Analyzing customer reviews with `GENERATE_TEXT` for sentiment and topic extraction.

Module 3: Prompt Engineering for BigQuery AI Functions

- Principles of effective prompt design for large language models.
- Structuring prompts within BigQuery SQL context.
- Techniques for few-shot prompting and context injection.
- Managing token limits and response parsing in BigQuery.
- Error handling and retry strategies for AI functions.
- Hands-on Lab: Optimizing product description generation and refinement using prompt engineering.

Module 4: BigQuery ML Remote Models and Inference

- Creating and managing BigQuery ML remote models for Gemini: `CREATE MODEL` syntax.
- Specifying model parameters and connection details.
- Performing batch inference using `ML.GENERATE_TEXT` and `ML.GENERATE_EMBEDDING`.
- Monitoring model usage and performance within BigQuery.
- Integrating with other BigQuery ML capabilities.
- Hands-on Lab: Building a news article summarization pipeline using a BigQuery ML remote model.

Module 5: Advanced Use Cases and Solutions

- Data enrichment: Generating missing descriptions, tags, or metadata for datasets.
- Semantic search and recommendation systems using vector embeddings.
- Anomaly detection leveraging embeddings for unusual data patterns.
- Automated code generation (e.g., SQL queries, Python scripts) from natural language prompts.
- Knowledge graph creation and entity extraction from unstructured text.
- Demo: Generating marketing copy and ad headlines from product catalog data.

Module 6: Responsible AI and Governance with Gemini in BigQuery

- Introduction to responsible AI principles for generative models.
- Identifying and mitigating bias in AI-generated content.
- Data privacy, security, and compliance considerations when using Gemini with sensitive data.
- Implementing content moderation and safety filters.
- Monitoring and auditing Gemini AI function usage for ethical guidelines.
- Hands-on Lab: Implementing guardrails and safety filters for user-generated content summarization.

Module 7: Integrating Gemini in BigQuery into Data Pipelines

- Orchestrating Gemini AI functions with BigQuery Scripting and stored procedures.
- Automating BigQuery ML remote model inference with Cloud Workflows or Cloud Composer.
- Real-time integration patterns using Cloud Functions or Cloud Run for AI- powered APIs.
- Leveraging Gemini-generated insights in Looker and other Business Intelligence tools.
- Data lineage and metadata management for AI-augmented data.
- Hands-on Lab: Building an automated data classification pipeline for support tickets using Gemini.

Module 8: Performance, Cost Optimization, and Future Directions

- Best practices for optimizing query performance with Gemini AI functions.
- Strategies for cost management and quota monitoring.

- Scaling Gemini in BigQuery solutions for large datasets and high throughput.
- Deployment considerations for production environments.
- Future trends: advancements in Gemini models, new BigQuery ML features, and ecosystem integrations.
- Course summary, further resources, and next steps for practitioners.

Your course overview ends here.
But your cloud journey doesn't!

Unlock advanced training and certification pathways with **CloudThat**.



About CloudThat

Established in 2012, CloudThat is an award-winning company and the first in India to offer cloud training and consulting services for individuals and enterprises worldwide. Recently, it won Google Cloud's New Training Partner of the Year Award for 2025, becoming the first company in the world in 2025 to hold awards from all three major cloud giants: AWS, Microsoft, and Google. CloudThat notably won consecutive AWS Training Partner of the Year (APJ) awards in 2023 and 2024 and the Microsoft Training Services Partner of the Year Award in 2024, bringing its total award count to an impressive 12 awards in the last 8 years.

CloudThat specializes in Cloud Migration, Data Platforms, DevOps, IoT, and cutting-edge technologies like Gen AI & AI/ML. As a Microsoft Solutions Partner, AWS Premier Tier Training Partner, Google Cloud Platform Partner, and collaborator with leading organizations like HPE, Oracle, Nvidia and Databricks, CloudThat has trained over 850,000 professionals across 600+ cloud certifications, empowering students and professionals worldwide to advance their skills and careers.

To know more about our cloud certification training, email at sales@cloudthat.com or call us at [+91-8880002200](tel:+91-8880002200) or [+1-8555588830](tel:+1-8555588830) for those from outside India.