Oracle AI Vector Search Deep Dive ELS

Oracle Cloud Infrastructure

DURATION	MODULES	COURSE CODE
2 Days	**10 Lectures**	**—**

Course Overview

Discover how Oracle Database 23aitransforms natural language questions into secure, actionable insights directly from your data.Master the integration of OCI Generative AI with Autonomous Database to unlock advanced AI-driven capabilities for your organization. You'll also learn to use natural language queries and vector search to streamline data exploration and decision-making processes.By the end of this course,you'll build expertise in retrieval-augmented generation (RAG) and embedding models, enhancing your ability to implement AI solutions. You'll also be able to equip yourself with future-proof skills in AI-powered data management, making you a valuable asset in the evolving tech landscape.

What You Will Learn

Module 1: Course Overview

- Course Overview
- Agenda
- Target Audience
- Learning Objectives

Module 2: Refresher on Vectors and Vector Embeddings

- Refresher: Vector & Vector Embeddings
- Vector Data Type
- Vector Embeddings
- Generating Vector Embeddings

Module 3: Exact Similarity Search

- Refresher: Exact Similarity Search
- Similarity Search Concepts
- Exact Similarity Search Techniques
- Euclidean and Euclidean Squared Distances
- Cosine Similarity

- Dot Product Similarity
- Manhattan Distance
- Hamming Similarity
- Vector Distance Operand

Module 4: Approximate & Multi-Vector Similarity Search

- Refresher: Approximate Similarity Search
- Multi-Vector Similarity Search
- Comparison of Exact vs Approximate Search

Module 5: Vector Index Fundamentals

- Refresher: Vector Index
- Vector Indexes Overview
- Why Vector Indexes are Needed
- Vector Pool
- In-Memory Neighbor Graph Vector Index (HNSW)
- Neighbor Partition Vector Index (IVF)
- Creating a Basic Vector Index
- Important Parameters & Limitations
- Using Vector Indexes
- Monitoring Index Accuracy
- Best Practices

Module 6: Retrieval-Augmented Generation (RAG) with Vector Search

- RAG Overview
- Vector Data Workflow for RAG
- RAG Workflow Step-by-Step
- Interacting with LLMs via RAG
- RAG with Oracle AI Vector Search

Module 7: Using Embedding Models with Oracle AI Vector Search

- Embedding Models Overview
- Using VECTOR_EMBEDDING() Function
- Creating and Vectorizing Tables
- Performing Similarity Searches
- Changing Embedding Models

Module 8: RAG with OCI Generative AI

- Overview of OCI Gen Service Integration
- RAG Steps:
- Text Extraction and Preparation
- Embedding Models and Vector Generation
- Similarity Search and Response Generation

- Building the Prompt
- Invoking the Chain
- Using PL/SQL for RAG
- Using Python for RAG

Module 9: Oracle AI Vector Search Supporting Features

- Exadata AI Storage Overview
- Refresher on GoldenGate Use Cases
- Oracle GoldenGate for Distributed AI & Vector Replication
- Real-Time Vector Hub for Generative AI
- Actionable AI/ML from Streaming Pipelines
- SQL Loader and Data Pump

Module 10: Course Summary

- Summary of Learning Outcomes
- Key Takeaways