Course Overview

Hello, and welcome to the **Vector Databases for RAG: An Introduction** course.

Gain expertise in using vector databases and improve your data retrieval skills in this hands-on course! Start by exploring the fundamental principles of similarity search and vector databases, understanding how they differ from traditional databases, and recognizing their importance in recommendation systems and Retrieval-Augmented Generation (RAG) applications.

Build on this foundation with practical experience using Chroma DB—one of the leading vector database solutions. Through interactive labs, learn to create collections, manage embeddings, and perform similarity searches with real-world datasets. Understand key concepts such as vector operations and database architecture to develop a strong grasp of Chroma DB's functionality.

Apply what you've learned by creating a real-world recommendation system powered by Chroma DB and an embedding model from Hugging Face. This project will enhance your understanding of how vector databases improve search and retrieval in AI-driven applications.

Throughout the course, engage in practical exercises to refine your skills in database management, similarity search techniques, and advanced vector operations. By the end, you'll have a thorough understanding of vector databases and similarity searches, which are essential components of the RAG pipeline.

This course is part of the [IBM RAG and Agentic AI Professional Certificate](#) ↗. By completing the program, you'll understand how to differentiate between vector databases and traditional databases based on their functionality and use cases. You will be able to perform fundamental database operations in Chroma DB, including updating, deleting, and managing collections. You will develop a thorough and comprehensive understanding of key internal mechanisms within RAG. You will also learn about similarity search techniques and apply them manually and with Chroma DB, and you will be able to develop recommendation systems using these techniques.

Prerequisites

This course is accessible to a wide range of professionals, but having some knowledge and experience of the Python programming language is required. An understanding of software applications and business systems and some familiarity with databases is advantageous but not required.

Objectives

After completing this course, you will be able to: