CORSEN AI Academy

Embeddings, RAG, and Semantic Search

Course syllabus

Version 2026-11

Overview

Build semantic search systems: text and code embeddings, RAG from scratch with FAISS, and advanced techniques (Corrective, Self-RAG, Adaptive).

Learning objectives

- Generate embeddings with mistral-embed and measure semantic similarity between two texts
- Split and batch a large corpus to vectorize it without exceeding API limits
- Index code with codestral-embed to enable semantic search in a repository
- Build a complete RAG pipeline: preparation, vectorization, FAISS storage, retrieval, and generation
- Compare Corrective RAG, Self-RAG, and Adaptive RAG to improve response relevance

Audience and prerequisites

Developers, engineers, and integrators looking to leverage Mistral AI via API in their applications.

Ability to read Python and JSON, command line basics. A Mistral AI Studio API key to try the examples.

Duration and format

Estimated duration : 2 h. 100% online self-paced course, accessible 24/7 from a browser. Progress saved, resume anytime. Estimated duration: 2 h.

Content

- Text Embeddings (4 lessons)
- Code Embeddings (4 lessons)
- RAG from Scratch (6 lessons)
- Advanced RAG (4 lessons)

Assessment

Self-assessment based on the Key takeaways concluding each lesson, and direct implementation of examples on your own examples. A final quiz of 5 corrected questions, at the end of the last lesson, validates acquired skills.

Accessibility

Course accessible to people with disabilities: text content compatible with screen readers, keyboard navigation, and compliant contrasts. For any adaptation needs, email hello@corsen.ai.

Price and recognition

Free

Certificate of completion issued by Corsen AI.

Access time

Immediate after account creation and email confirmation.

Corsen AI — Paris, France — SIRET 10297040700015
academy@corsen.ai — https://academy.corsen.ai