Embeddings, RAG et Recherche Sémantique

Programme de formation

Version 2026-11

Présentation

Embeddings OpenAI, bases vectorielles, recherche sémantique, RAG complet et applications en production.

Objectifs pédagogiques

- Générer des embeddings et mesurer la proximité sémantique entre textes par similarité cosinus
- Indexer un corpus dans une base vectorielle et mettre en place une recherche hybride sémantique et lexicale
- Construire un pipeline RAG complet : chunking, récupération, reranking puis génération
- Évaluer la pertinence d'un pipeline RAG et arbitrer entre File Search et une architecture sur mesure
- Exploiter les embeddings pour classifier, regrouper, recommander et détecter anomalies et doublons

Public visé et prérequis

Développeurs, ingénieurs et intégrateurs souhaitant exploiter ChatGPT par API dans leurs applications.

Savoir lire du Python et du JSON, bases de la ligne de commande. Une clé d'API de la plateforme OpenAI pour essayer les exemples.

Durée et modalités

Durée estimée : 2h15. Formation 100 % en ligne, en autonomie, accessible 24h/24 depuis un navigateur. Progression enregistrée, reprise possible à tout moment. Durée estimée : 2h15.

Contenu

- Embeddings OpenAI (5 leçons)
- Recherche sémantique (5 leçons)
- RAG — Retrieval-Augmented Generation (6 leçons)
- Applications (6 leçons)

Modalités d'évaluation

Auto-évaluation à partir des « Points clés à retenir » qui concluent chaque leçon, et mise en pratique directe des exemples sur vos propres cas. Un quiz final de 5 questions corrigées, en fin de dernière leçon, permet de valider les acquis.

Accessibilité

Formation accessible aux personnes en situation de handicap : contenus textuels compatibles avec les lecteurs d'écran, navigation au clavier et contrastes conformes. Pour tout besoin d'adaptation, écrivez à hello@corsen.ai.

Tarif et reconnaissance

Gratuit

Attestation de fin de formation délivrée par Corsen AI (non éligible au CPF à ce jour).

Délai d'accès

Immédiat après la création du compte et la validation de l'adresse email.

Corsen AI — Paris, France — SIRET 10297040700015
Organisme de formation en cours de déclaration d'activité.
academy@corsen.ai — https://academy.corsen.ai