

Mis à jour le 27/06/2025

S'inscrire

Formation ChromaDB

2 jours (14 heures)

Présentation

Maîtrisez ChromaDB pour construire des moteurs de recherche sémantique puissants et des bases vectorielles prêtes à intégrer vos applications IA. Cette formation vous accompagne pas à pas dans la conception, l'indexation et l'interrogation de données vectorielles.

Vous apprendrez à générer et manipuler des embeddings à partir de textes, à les indexer efficacement dans Chroma, et à optimiser vos collections pour des recherches rapides, précises et filtrées par métadonnées.

Vous saurez créer des collections persistantes, interroger vos données avec des critères avancés, visualiser les distances sémantiques entre vos contenus, et maintenir une base vectorielle adaptée à la production.

Vous découvrirez également comment intégrer Chroma dans un système RAG complet via LangChain, en combinant retrieval sémantique et génération de réponse contextuelle avec un LLM comme GPT. Vous serez ainsi en mesure de créer vos propres assistants intelligents connectés à votre base de connaissances.

Comme pour toutes nos formations, elle se déroulera sur la toute dernière version de l'outil : [ChromaDB](#)

Objectifs

- Comprendre les principes de la recherche vectorielle, le rôle des embeddings, et les fondements de la similarité sémantique dans les systèmes d'IA modernes.
- Installer, configurer et administrer une base vectorielle ChromaDB en mode mémoire ou persistant, avec gestion fine des collections, documents et métadonnées.
- Générer, indexer et interroger efficacement des embeddings de texte à l'aide de modèles pré-entraînés (OpenAI, SentenceTransformers), avec des requêtes filtrées et précises.
- Concevoir des workflows complets d'ingestion, de découpage, d'indexation et d'interrogation de données textuelles dans une base Chroma à des fins de recherche intelligente.

- Exploiter ChromaDB dans une architecture RAG en l'intégrant à LangChain pour construire des assistants conversationnels connectés à une base documentaire vectorisée.
- Superviser, sécuriser et optimiser les bases vectorielles en production, en abordant les problématiques de performance, de gouvernance, de persistabilité et de mise à l'échelle.

Public visé

- Développeurs
- Data Scientists
- Chercheur IA

Pré-requis

- Connaissances de base sur les APIs REST
- Connaissance dans un langage de développement moderne Python

Programme de la formation ChromaDB

Introduction à la recherche vectorielle

- Représentation sémantique des textes
- Embeddings courants : OpenAI, HuggingFace, BERT
- Cosine similarity, Euclidean, Manhattan
- Notions de top-k, seuils de score

Architecture et fonctionnement de ChromaDB

- Collections, Documents, Embeddings, Metadata
- Index interne (FAISS-like, local storage)
- Mémoire uniquement (:memory:)
- Persistante (dossier local)
- Documents texte, paires clé/valeur, métadonnées optionnelles

Mise en place de ChromaDB

- pip install chromadb
- Dépendances utiles (tqdm, langchain pour test)
- chromadb.Client()
- Dossier de stockage (persist_directory)
- create_collection(), nommage, duplication, gestion
- Ajout de documents et embeddings

Indexation et interrogation des données

- Embeddings via OpenAI ou sentence-transformers
- add() et upsert() : quand utiliser quoi
- query(query_embeddings=..., n_results=...)
- Filtrage par métadonnées
- delete(), modify(), reset() : gérer ses données
- Export / import éventuel (JSON, backup local)

Visualisation et analyse des résultats

- ids, documents, metadatas, distances
- Utiliser pandas ou prettytable pour inspection rapide
- t-SNE, PCA pour représentation 2D

Cas pratiques d'utilisation

- Embedding des phrases de transcript
- Récupération de passages pertinents
- PDFs convertis en chunks indexés
- Affichage des résultats avec contexte
- Utilisation des métadonnées

Intégration avec LangChain

- Embedding ? Vector Store ? Retriever ? LLM
- Chroma.from_documents(documents, embedding)
- as_retriever(), similarity_search()
- Utilisation de RetrievalQA
- Injection du contexte dans le prompt

Bonnes pratiques & Ressources

- Documentation officielle Chroma
- Chunking optimal, stockage persistant, embeddings légers
- Projets d'approfondissement
 - Chatbot juridique
 - moteur de recherche RH
 - assistant support client

Sociétés concernées

Cette formation s'adresse à la fois aux particuliers ainsi qu'aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée ou bien à

acquérir des connaissances métiers spécifiques ou des méthodes modernes.

Positionnement à l'entrée en formation

Le positionnement à l'entrée en formation respecte les critères qualité Qualiopi. Dès son inscription définitive, l'apprenant reçoit un questionnaire d'auto-évaluation nous permettant d'apprécier son niveau estimé sur différents types de technologies, ses attentes et objectifs personnels quant à la formation à venir, dans les limites imposées par le format sélectionné. Ce questionnaire nous permet également d'anticiper certaines difficultés de connexion ou de sécurité interne en entreprise (intraentreprise ou classe virtuelle) qui pourraient être problématiques pour le suivi et le bon déroulement de la session de formation.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiples permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.