

Mis à jour le 18/10/2024

S'inscrire

Formation LanceDB

2 jours (14 heures)

Présentation

LanceDB est une base de données conçue pour l'IA spécialisée dans la gestion et la recherche de données vectorielles, idéale pour les applications d'intelligence artificielle, de machine learning et de recherche sémantique. Optimisée pour les embeddings, elle permet des recherches rapides et précises sur des données complexes, notamment des images, textes et autres types de contenu.

Dans ce programme, vous apprendrez à installer et configurer LanceDB, à manipuler l'interface utilisateur et à maîtriser les commandes de base. Nous explorerons les avantages de LanceDB par rapport à d'autres bases de données et vous verrez comment l'utiliser dans divers scénarios comme la recherche vectorielle et le traitement des données.

Nous aborderons des concepts fondamentaux tels que l'indexation vectorielle et scalaire, la gestion des embeddings pour la recherche sémantique, ainsi que l'intégration avec des outils comme Pandas et DuckDB. Vous apprendrez aussi à créer des tables et à construire des index optimisés pour des performances accrues, y compris la recherche hybride.

À la fin de cette formation, vous serez capable de migrer des données vers LanceDB, d'optimiser la récupération d'informations à grande échelle et de résoudre des problèmes de performance. Vous maîtriserez l'intégration des modèles de machine learning pour des recherches plus intelligentes, vous dotant d'une expertise pratique en recherche vectorielle et sémantique.

Cette formation vous sera présentée avec les [dernières ressources](#) de la technologie.

Objectifs

- Installer et configurer efficacement LanceDB
- Maîtriser l'interface utilisateur et les commandes de base
- Comprendre et appliquer l'indexation vectorielle et scalaire

- Intégrer des modèles personnalisés et utiliser des embeddings
- Optimiser les performances et gérer la migration des données

Public visé

- Développeurs
- Data Scientists
- Ingénieurs

Pré-requis

- Une bonne compréhension des bases de données relationnelles et non relationnelles
- Familiarité avec le concept d'embeddings et la recherche vectorielle
- Expérience dans l'intégration de modèles de machine learning et dans l'optimisation de bases de données

PROGRAMME DE NOTRE FORMATION LanceDB

INTRODUCTION À LanceDB

- Présentation de LanceDB
- Installation et configuration initiale
- Aperçu de l'interface utilisateur et des commandes de base
- Avantages de LanceDB par rapport à d'autres bases de données
- Scénarios d'utilisation typiques et cas d'études

CONCEPTS FONDAMENTAUX DE LanceDB

- Introduction à la recherche vectorielle et son importance
- Comprendre l'indexation : méthodes et meilleures pratiques
- Options de stockage et implications sur les performances
- Gestion des données : insertion, mise à jour, suppression
- Exploration des types d'index : vectoriel, scalaire, et hybride

PRATIQUE AVEC LES TABLES ET LES INDEX

- Création et manipulation de tables dans LanceDB
- Construction et optimisation d'un index vectoriel
- Recherche vectorielle approfondie et recherche en texte intégral
- Mise en place d'un index scalaire et recherche correspondante
- Utilisation de la recherche hybride pour des résultats optimisés

GESTION DES EMBEDDINGS ET INTÉGRATIONS

- Comprendre les embeddings et leur rôle clé dans les recherches sémantiques
- Utilisation des fonctions d'embeddings et intégration de modèles personnalisés
- Exemples pratiques : recherche sémantique multilingue et embeddings multimodaux CLIP
- Intégration de LanceDB avec Pandas, PyArrow, Polars, et DuckDB
- Démonstration d'intégrations avancées comme LangChain et LlamaIndex

MIGRATION ET OPTIMISATION DES PERFORMANCES

- Guide de migration : stratégies pour transférer des données vers LanceDB
- Identification des défis courants et solutions lors de la migration
- Techniques avancées pour le tuning des performances de récupération des données
- Ajustement des paramètres de stockage pour améliorer les performances
- Cas pratiques et résolution de problèmes en direct

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.

