Vektor Datenbanken

[Untertitel des Dokuments]

HAW-Hamburg, Berliner Tor 21

Kurzdokumentation für Vektordatenbanken

# Einleitung

Vektor Datenbanken werden genutzt, um Daten in eine numerische Darstellung zu konvertieren. Das ermöglicht einen schnelleren Zugriff auf viele Daten.

# Abstrakt

Aufgrund der Komplexität von großen Datensätzen sind Vektordatenbanken für Maschinenlernen oder KI-Tätigkeiten unerlässlich, damit die Verarbeitung von Anfragen oder Berechnungen möglichst schnell und effizient abgearbeitet werden können. [1]

# Funktionsweise von Vektordatenbanken

Vektordatenbanken speichern Informationen numerisch in einem n-Dimensionalen Raum als Datenpunkte. Diese Daten sind in diesem Vektorraum mit Methoden wie KD-Bäumen oder Locality-Sensitive-Hashing (LSH) indiziert, um eine effiziente Ähnlichkeitssuche zu gewährleisten. [1] Abfragen werden mit indizierten Vektoren verglichen, indem Entfernungsmetriken wie euklidische Distanz oder Kosinusähnlichkeit verwendet werden, die für Anwendungen wie Empfehlungssysteme entscheidend sind. Die ähnlichsten Vektoren werden schnell abgerufen und ermöglichen Echtzeitanwendungen wie Suchmaschinen und KI-DR.

## Beispiele

# Begriffe

## KD-Bäume

Baum mit Parent-Child Verbindung in n Richtungen. [2]

## Locality-Sensitive-Hashing

Clustering von verschiedenen Informationen, die ähnliche Bedeutungen haben. [3]

## Euklidische Distanz

Vektordistanz über Pythagoras. [4]

# Quellen

[1] Top 15 Vector Databases that You Must Try in 2025  
<https://www.geeksforgeeks.org/top-vector-databases/>  
(28.04.2025, 16:28)

[2] KD-Bäume  
<https://de.wikipedia.org/wiki/K-d-Baum>  
(28.04.2025, 17:07)

[3] Locality-Sensitive-Hashing  
<https://en.wikipedia.org/wiki/Locality-sensitive_hashing>  
(28.04.2025, 17:16)

[4] Euklidische Distanz  
<https://de.wikipedia.org/wiki/Euklidischer_Abstand>  
(28.04.2025, 17:26)

[5] Embeddings and Vector Databases With ChromaDB  
<https://realpython.com/chromadb-vector-database/>  
(29.04.2025, 13:54)