Vector Databases Versus Traditional Databases

Estimated time needed: **10** minutes

Welcome to this reading, where you can compare vector databases with traditional databases.

What you will learn

After completing this reading, you will be able to:

- Explain the concept of a vector database.
- Summarize how vector databases differ from traditional relational databases.
- Describe how vector databases and relational databases store data.
- Explain the difference between a vector library and a vector database.
- Explain the key differences in how vector and relational databases manage data, including data presentation, searching and retrieval, indexing, scalability, and typical applications.
- Evaluate the advantages and disadvantages of using vector databases versus relational databases based on specific data management needs.
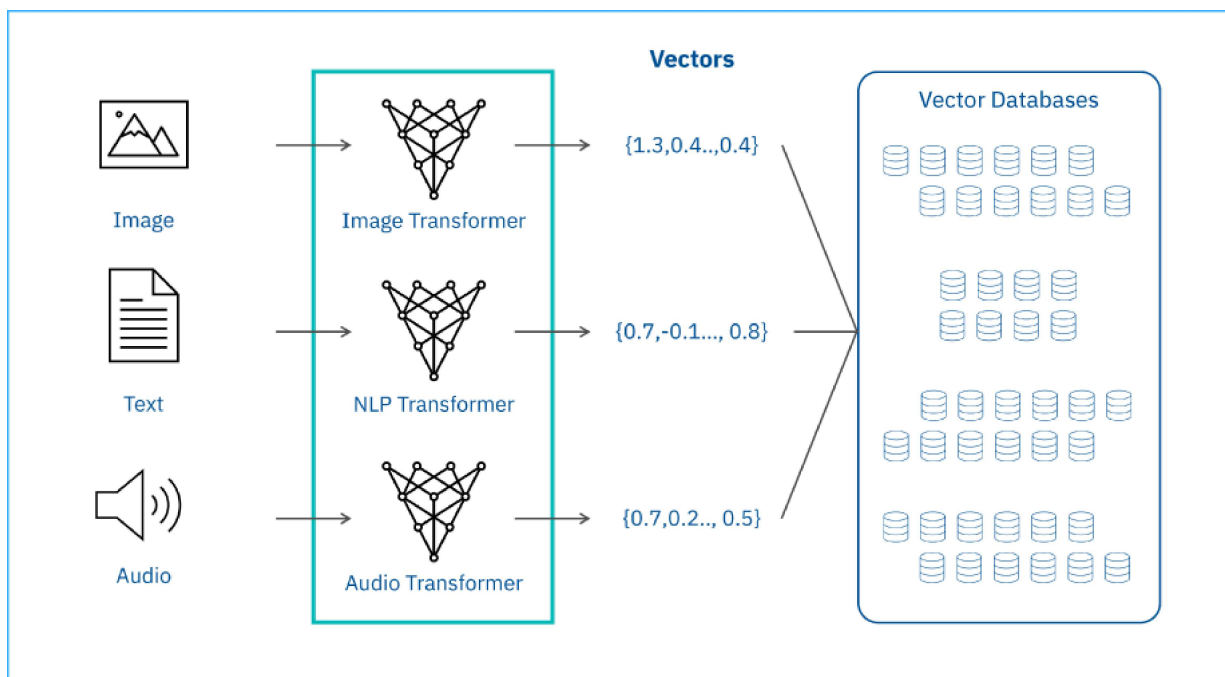
Vector databases

A vector database is a specialized database designed to store and query vectorized data rapidly. Unlike conventional databases, which organize data in tables, a vector database represents data as vectors in a multi-dimensional space. These vectors encapsulate essential attributes of the items they represent, making vector databases ideal for tasks requiring similarity searches, nearest neighbor queries, and assessing distances or similarities between vectors.

Vector databases data storage

Vector databases store data formatted so that numerical vectors depict each data item. These numbers signify various attributes or features of the object, with each vector dimension corresponding to a specific attribute. For instance, in an image database, each image may be represented as a vector of pixel values, while a text database might represent each piece of text as a vector of word frequencies.

In the following illustration you can see that you can take images, text (including text strings), and audio, then apply the appropriate transformer to create vector embeddings, and then store the data in vector databases for further use.

Vector libraries

Vector libraries within in-memory vector databases. Besides libraries, Vector databases use pre-configured algorithms to store and update data. You'll find commercially available vector libraries for similarity. In contrast, vector databases have full CRUD (create, read, update, and delete) capabilities. Additionally, vector databases are often a part of enterprise-level production deployments.
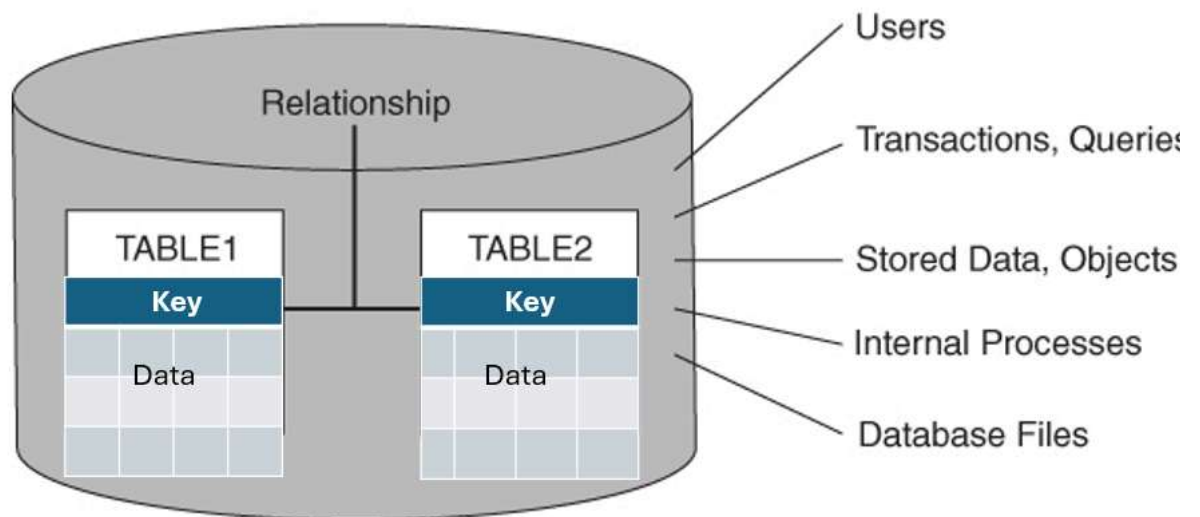
Relational database organization

A relational database organizes data into tables, utilizing rows and columns. Relational databases use structured query language (SQL) for data querying and manipulation, adhering to the relational model. Relational databases excel in managing structured data where relationships between different entities are well-defined.

Relational database data storage

Relational databases store data in tables, each representing a distinct entity or relationship.

Each row in a table corresponds to a record, while each column represents a property or attribute. Tables contain data and get connected, having relationships. With other tables using keys, such as main and foreign keys.

In a relational database, users can perform transactions, queries, and internal processes on stored data or objects (database files) within rows and columns of tables.

The database structures data in these tables and establishes connections between the tables using keys identified as primary and foreign keys. A relational database query manipulates the rows and columns using SELECT, INSERT, UPDATE, and DELETE operations.

Vector and relational databases compared

Check out this comparative table to understand the distinctions between vector and relational databases, which is crucial for selecting the appropriate database type for specific needs.

Function	Traditional databases	Vector databases
Data Representation	Traditional databases organize data in a structured format using tables, rows, and columns, ideal for relational data.	Vector databases represent data as multi-dimensional vectors, efficiently encoding complex and unstructured data like images, text, and sensor data.
Data Search and Retrieval	SQL queries are suited for traditional databases with structured data.	Vector databases specialize in similarity searches and retrieving vectorized data, facilitating tasks like image retrieval, recommendation systems, and anomaly detection.
Indexing	Traditional databases employ indexing methods like B-trees for efficient data retrieval.	Vector databases use indexing structures like metric trees and hashing suited for high-dimensional spaces, enhancing nearest-neighbor searches and similarity assessments.
Scalability	Scaling traditional databases can be challenging, often requiring resource augmentation or data sharding.	Vector databases are designed for scalability, especially in handling large datasets and similarity searches, using distributed architectures for horizontal scaling.
Applications	Traditional databases are pivotal in business applications and transactional systems where structured data is processed.	Vector databases shine in analyzing vast datasets, supporting fields like scientific research, natural language processing, and multimedia analysis.

Recap

In this reading, you learned that:

- A vector database is a specialized database representing data numerically as vectors in a multi-dimensional space.
- While traditional databases are adept at managing structured data and transactional operations, vector databases handle high-dimensional data and perform rapid similarity searches.
- Vector libraries read and update data; however, vector databases can perform create, read, update, and delete (CRUD) functions.
- Vector databases store data formatted so that numerical vectors depict each data item.
- Relational databases organize data into tables, utilizing rows and columns.
- Traditional databases use SQL queries, and vector databases use similarity search to retrieve data.

Author

Richa Arora