What is vector Database?

A vector database is a type of database that stores data as high-dimensional vectors, which are mathematical representations of features or attributes. Each vector has a certain number of dimensions, which can range from tens to thousands, depending on the complexity and granularity of the data. The vectors are usually generated by applying some kind of transformation or embedding function to the raw data, such as text, images, audio, video, and others. The embedding function can be based on various methods, such as machine learning models, word embeddings, feature extraction algorithms.

The main advantage of a vector database is that it allows for fast and accurate similarity search and retrieval of data based on their vector distance or similarity. This means that instead of using traditional methods of querying databases based on exact matches or predefined criteria, you can use a vector database to find the most similar or relevant data based on their semantic or contextual meaning.

For example, you can use a vector database to:

find images that are similar to a given image based on their visual content and style

find documents that are similar to a given document based on their topic and sentiment

find products that are similar to a given product based on their features and ratings