

This is a sample PDF document used for testing the Pinecone PDF chatbot.

The document contains information about how vector databases work. Vector databases are designed to store and search over high-dimensional vectors. They are commonly used in machine learning and AI applications, especially for semantic search.

Pinecone is a fully managed vector database service that allows fast and scalable similarity search. It supports real-time ingestion and querying of vector data.

To use Pinecone, you first embed your data into vectors using an embedding model like OpenAI or Sentence Transformers, then upsert those vectors to Pinecone. Queries are also embedded and matched using similarity metrics like cosine similarity.

This demo showcases how to extract text from PDFs, embed them, and search them using a chatbot.