

Aplicaciones LLM

Conceptos avanzados para optimizar la técnica RAG

Conceptos avanzados para optimizar la técnica RAG

- Técnicas alternativas para indexar y almacenar vectores.
- Técnicas alternativas para medir la similarity entre vectores.

Técnicas para almacenar e indexar vectores

- Hierarchical Navigable Small World, HNSW
- Locally Sensitive Hashing, LSH
- Product Quantization, PQ
- Filtering (pre and post filtering)
- Random Projections

Técnicas para medir la similarity entre vectores

- Las más comunes son:
 - cosine similarity (ángulo entre vectores).
 - euclidean distance (línea recta entre items).