Extra slides about pgvector and indexing

By Tibs (they / he)

Slides and source code at
https://github.com/aiven-labs/pgvector-find-faces-talk

tony.ibbs@aiven.io / *https://aiven.io/tibs* / *@much_of_a*

Indexing

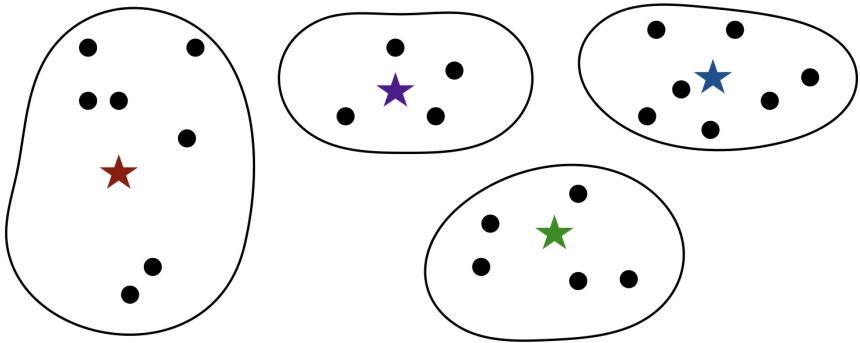
Speeds up the *use* of embeddings, but gives *approximate* results.

- IVFFlat: Inverted File with Flat Compression

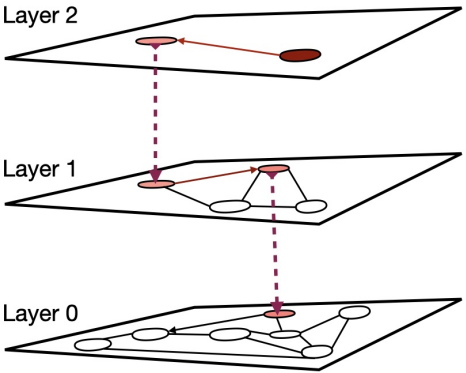
  Slower to search, quicker to build, smaller index, needs rebuilding
- HNSW: Hierarchical Navigable Small Worlds

  Faster to search, slower to build, adapts to new data

IVFFlat: Inverted File with Flat Compression



HNSW: Hierarchical Navigable Small Worlds



So which to choose?

Advice from Vector Indexes in Postgres using pgvector: IVFFlat vs HNSW

- If you care more about index size, then choose IVFFlat.
- If you care more about index build time, then select IVFFlat.
- If you care more about speed, then choose HNSW.
- If you expect vectors to be added or modified, then select HNSW.