Vector Search Introduction

Yinghao He

he923@purdue.edu

About me

- First year PhD student in Computer Science
- Graduated from Tongji University, Shanghai, China
- Working on building efficient vector databases

1. What is vector search?

2. Applications in LLM

3. Integrate vector search in MySQL

What is vector search?

Audio 1

Audio 2

What is vector search?

Traditional database

Animal Table

Content	Label
	A joyful cat
	An unpleasant cat
	An angry cat
	A peaceful cat
	A calm mouse
	A sleepy panda

Different Types of data in modern applications:

1. Images
2. Audio
3. Video
4. Biological data

.....

What is vector search?

Traditional database

Animal Table

Content	Label
	A joyful cat
	An unpleasant cat
	An angry cat
	A peaceful cat
	A calm mouse
	A sleepy panda

Meanings beyond labels

Happiness

Stable emotion

- Humans can easily understand these relationships
- For databases, do they understand?

What is vector search?

Traditional database

Animal Table

Content	Label
	A joyful cat
	An unpleasant cat
	An angry cat
	A peaceful cat
	A calm mouse
	A sleepy panda

Key words matching

Limitations of Traditional Search

1. **Keyword Dependency:** Relies strictly on exact string matches.
2. **Semantic Blindness:** Fails to recognize synonyms or context .
3. **Manually Label:** Requires human tagging for all data.
4. **Cross-Modal Barrier:** Cannot directly query audio or images.

What is vector search?

Convert data into vectors

✖ “Does this label contain word cute?”

✔ “Is this item semantically close to what the user want?”

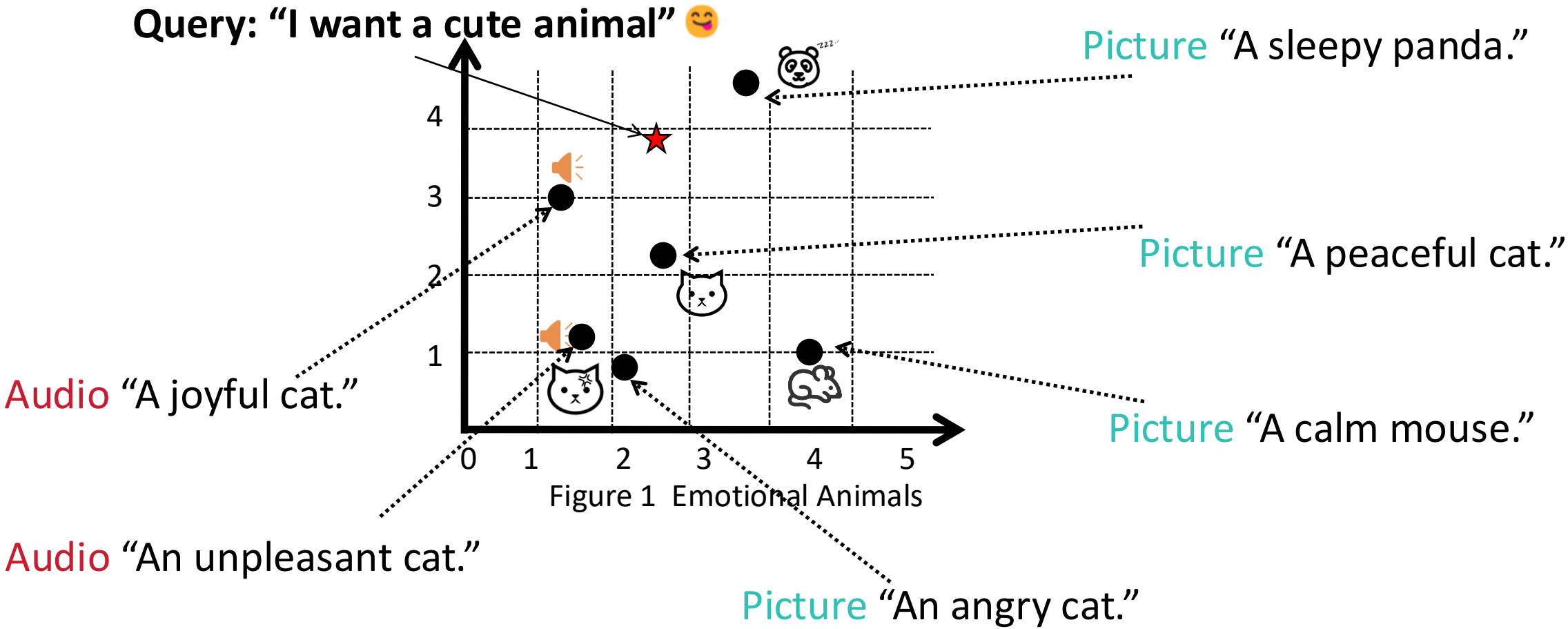


Figure 1 Emotional Animals

What is vector search?

1. Vector Search = Embeddings + Distance Metrics

Embedings: Map data to vectors in high-dimensional space. *Example*: “cat” -> [1.26, 2.04, -0.25, ...]

Distance Metrics: A way to measure similarity. (e.g., L2 distance, cosine similarity)
If two vectors are close, then their meanings are likely similar.

2. Search Algorithms: KNN vs ANN

KNN

* **Exhaustive search:** Compute distance with every vector and return the nearest ones

* **100% accuracy** for it checks all points

* **Very slow:** O(N) complexity, doesn’t scale well for large dataset(eg. **50K**)

* **For small dataset**

K-**N**earest **N**eighbors

ANN

* **Selective search**: explore only a subset of promising candidates

* **Much faster:** sub-linear complexity

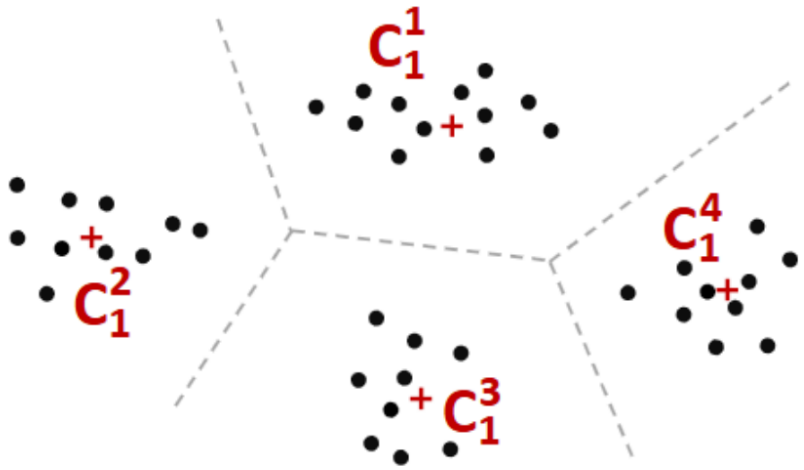
* **~95% accuracy:** tradeoff between speed and quality, it may not return the exact best answers, but it can still achieve high accuracy

* **For large dataset:** much more widely used

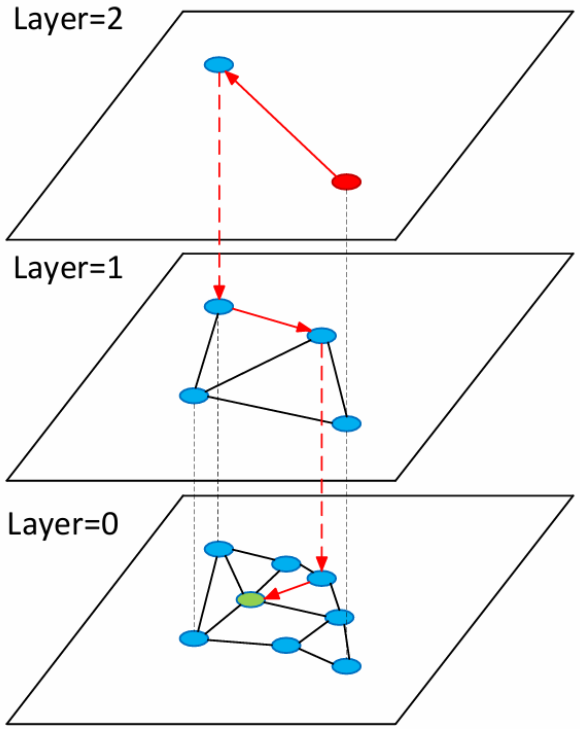
Approximate **N**earest **N**eighbors

What is vector search?

3. Major ANN Algorithms



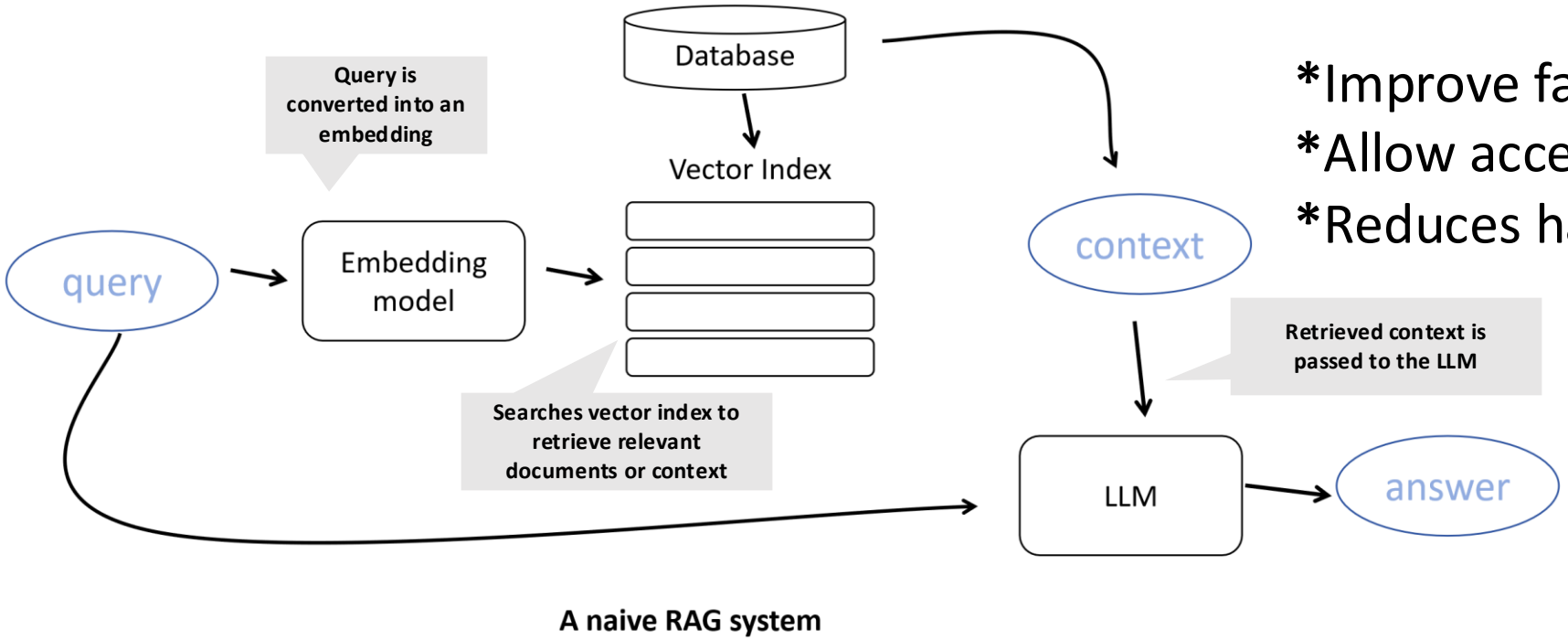
Clustering-based Algorithm
(eg. IVF, **I**n**v**erted **f**ile index)



Graph-based Algorithm
(eg. HNSW, **H**ierarchical **N**avigable **S**mall **W**orld graphs)

Applications in LLM

1. RAG(Retrieval-Augmented Generation)



A naive RAG system

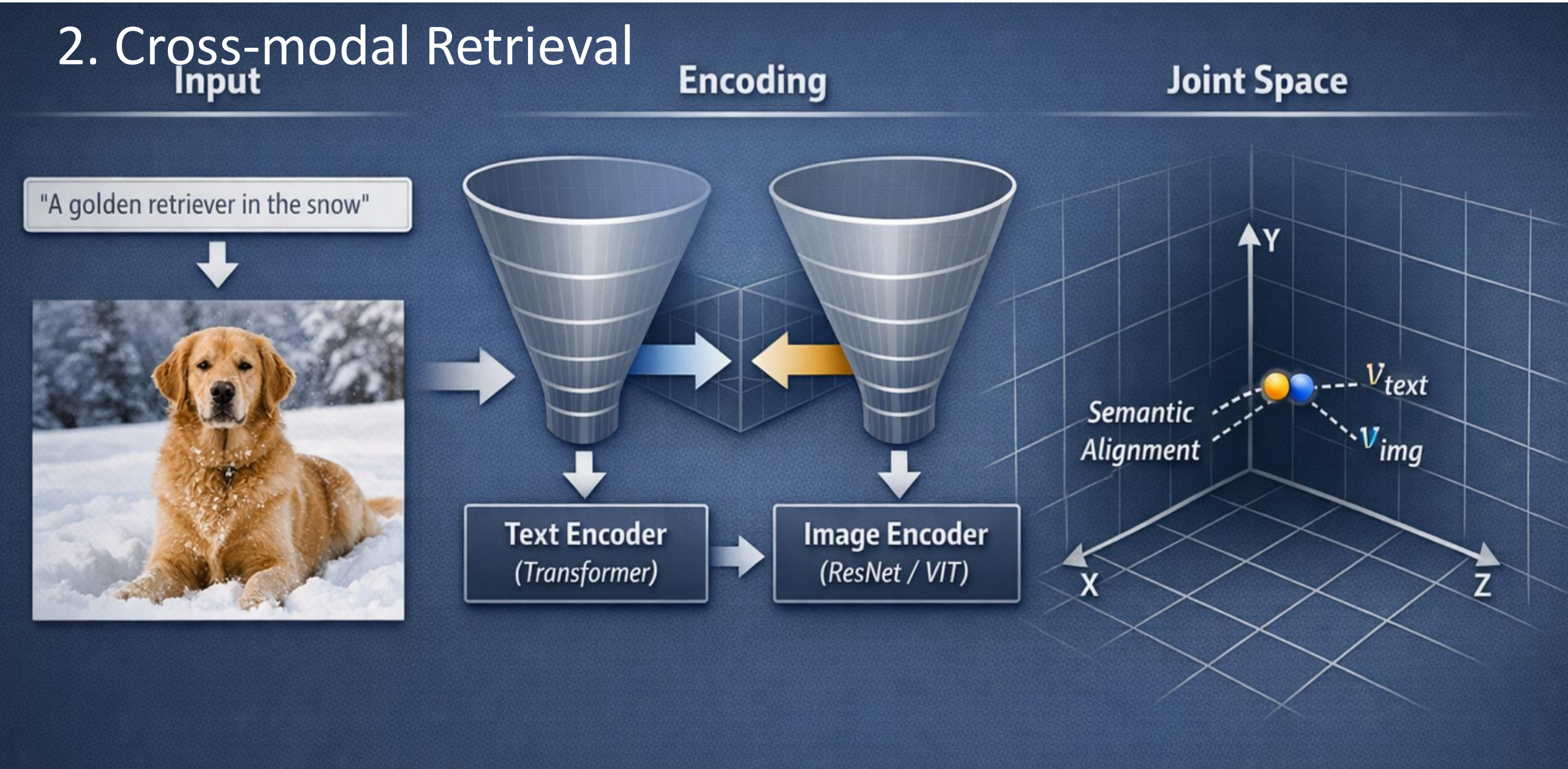
Benefits of RAG:

*Improve factual accuracy
*Allow access to up-to-date information
*Reduces hallucination

Act as the bridge between the user query and relevant external knowledge

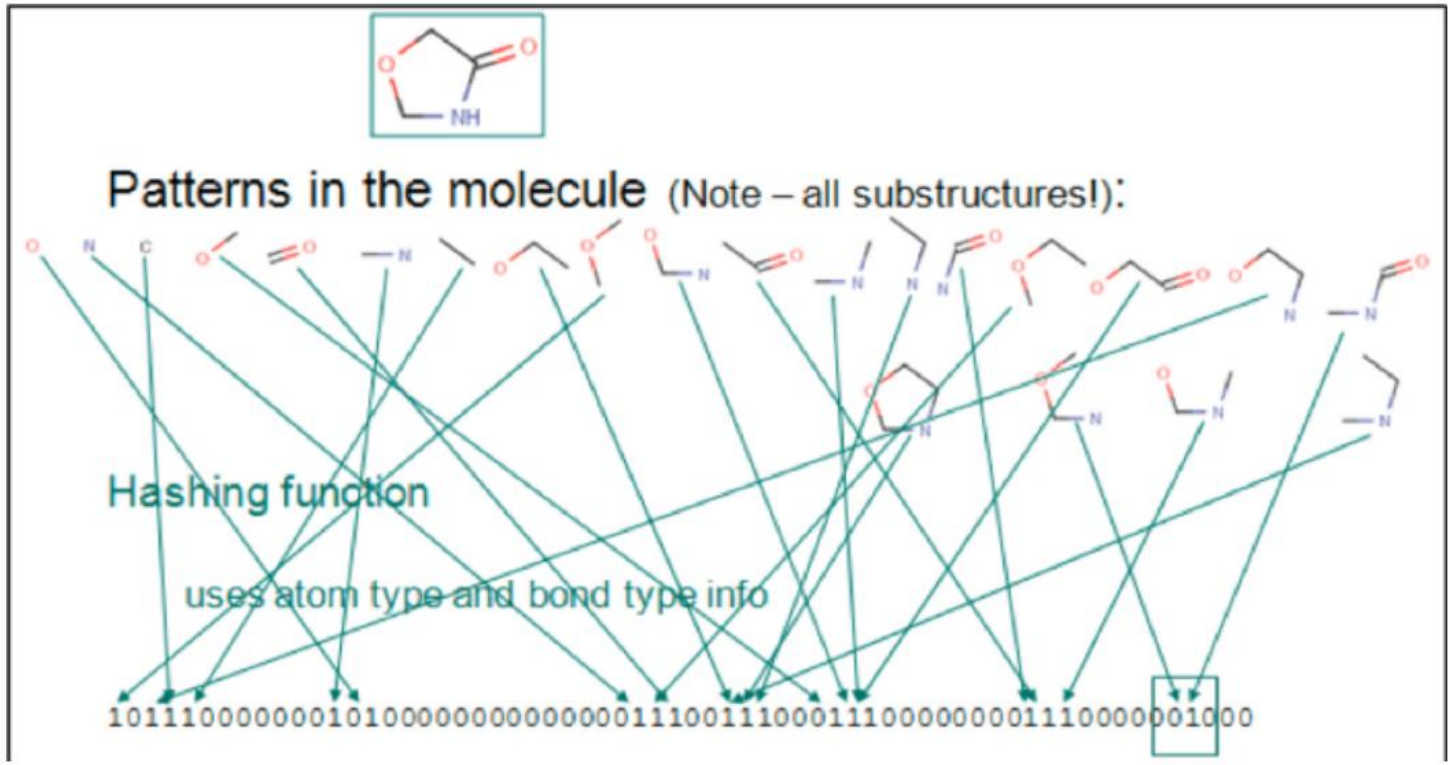
Applications in LLM

2. Cross-modal Retrieval



Applications in LLM

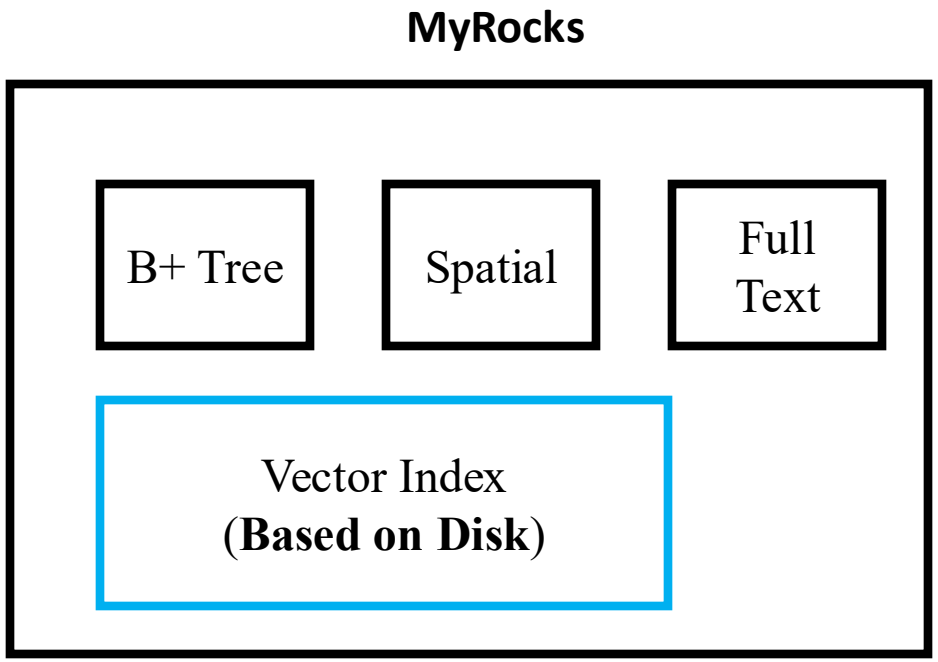
3. Bioinformatics



2-identifying-patterns-molecules.png

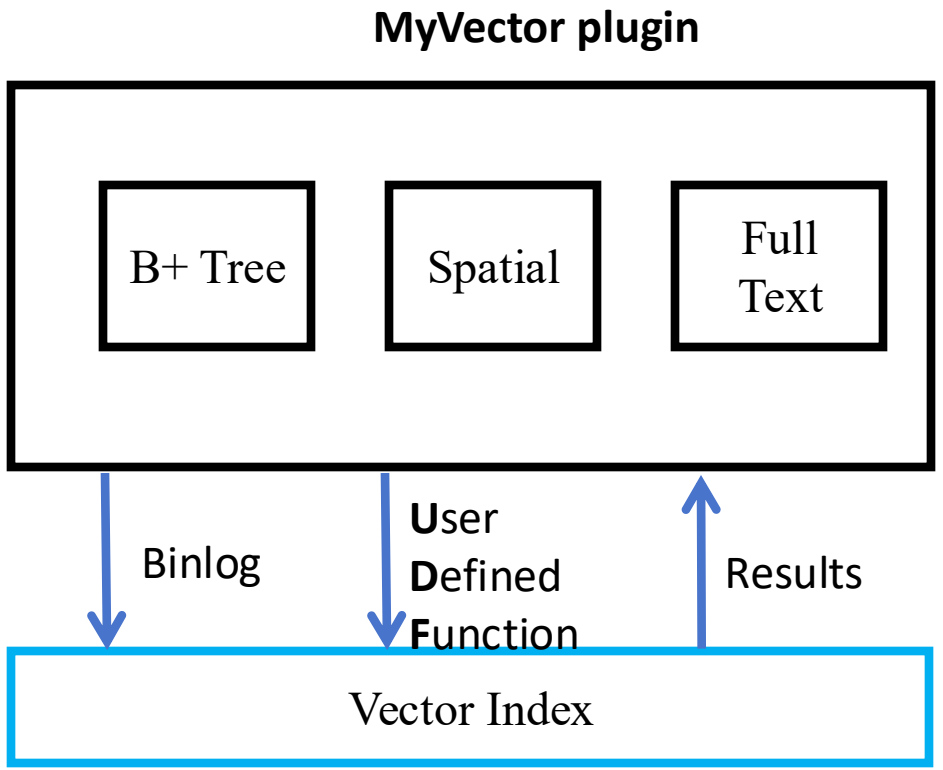
Vector Search in Bioinformatics

Intergrate vector search in MySQL



Pros: Vector index stored as tables
Easy to implement

Cons: Mainly disk-based, slower search performance



Pros: Treats vector index as a plugin
Runs in memory, so it is very fast

Cons: Lacks strong consistency

Intergrate vector search in MySQL

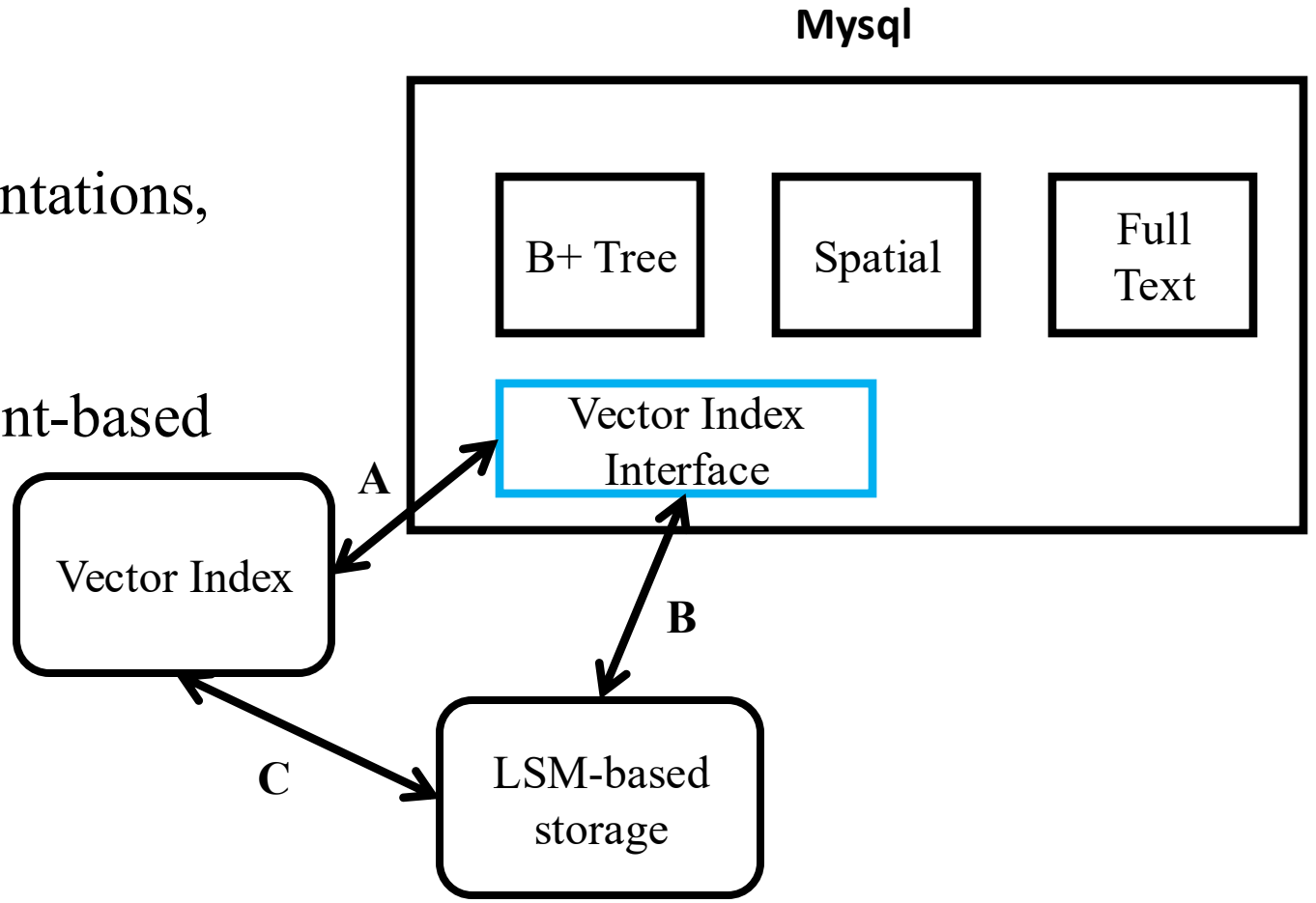
Vector Index Interface: Provides a standardized API for **CRUD** operations with guaranteed ACID compliance.

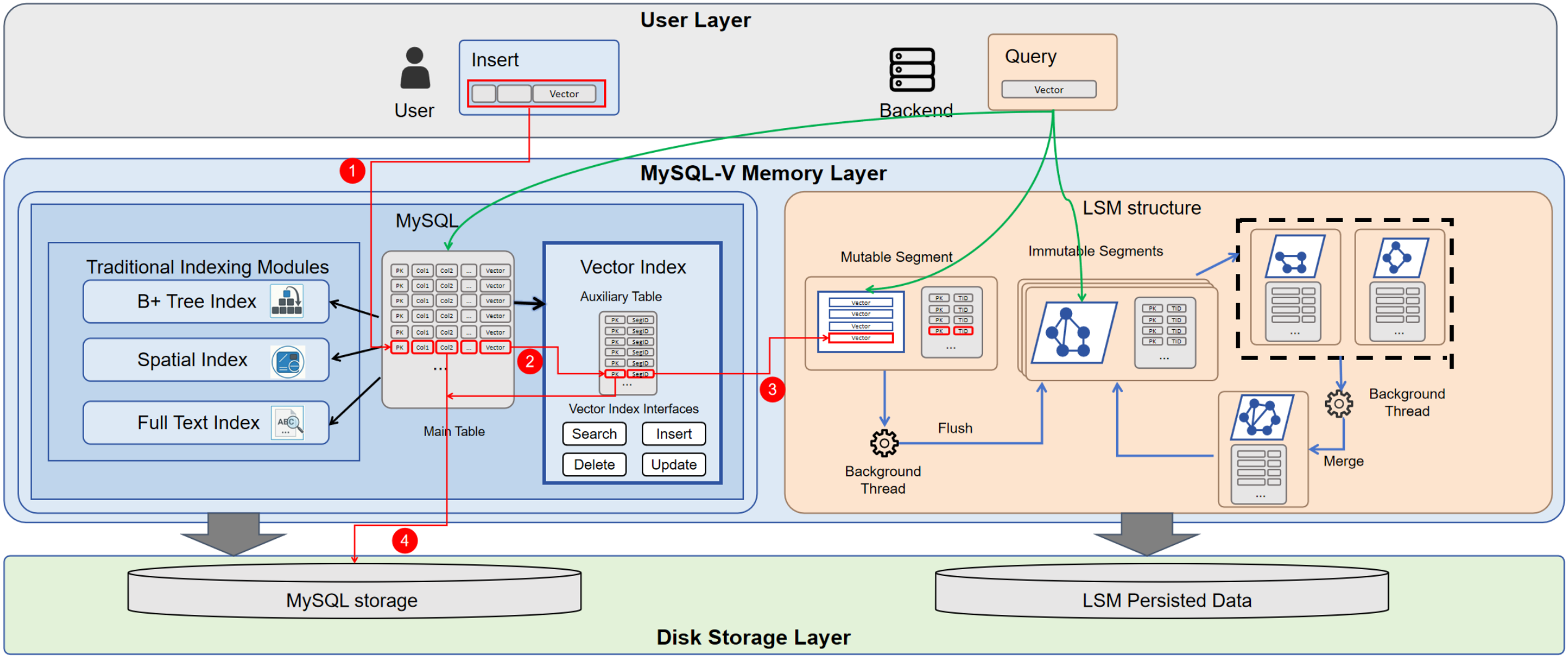
Vector Index: Supports diverse implementations, such as *HNSWlib/Faiss*/DiskANN....

LSM-based framwork: Manages segment-based storage for vector indices and background merge operations.

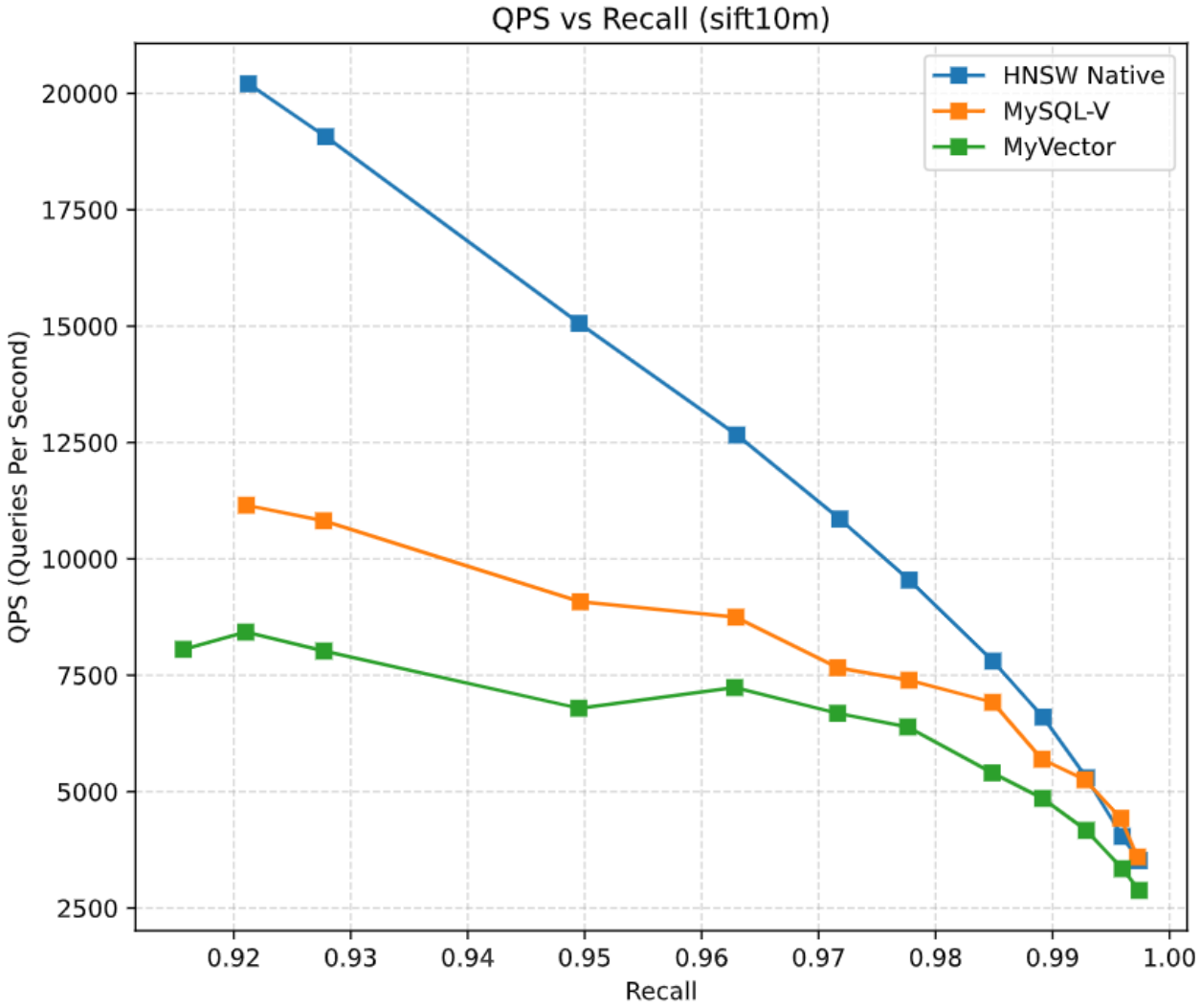
Pros: Strong consistency
Better performance

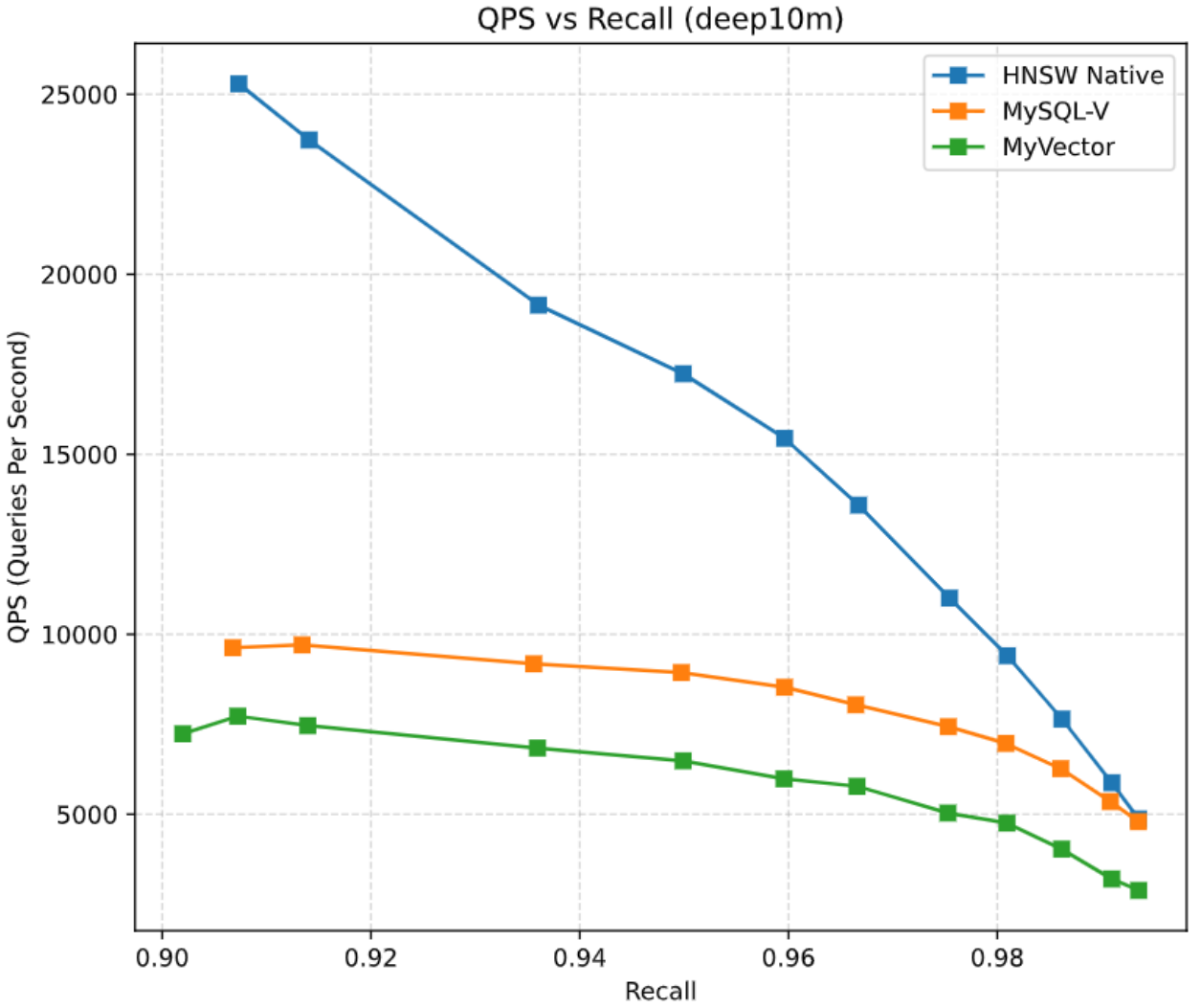
Cons: Complex code, but not for users



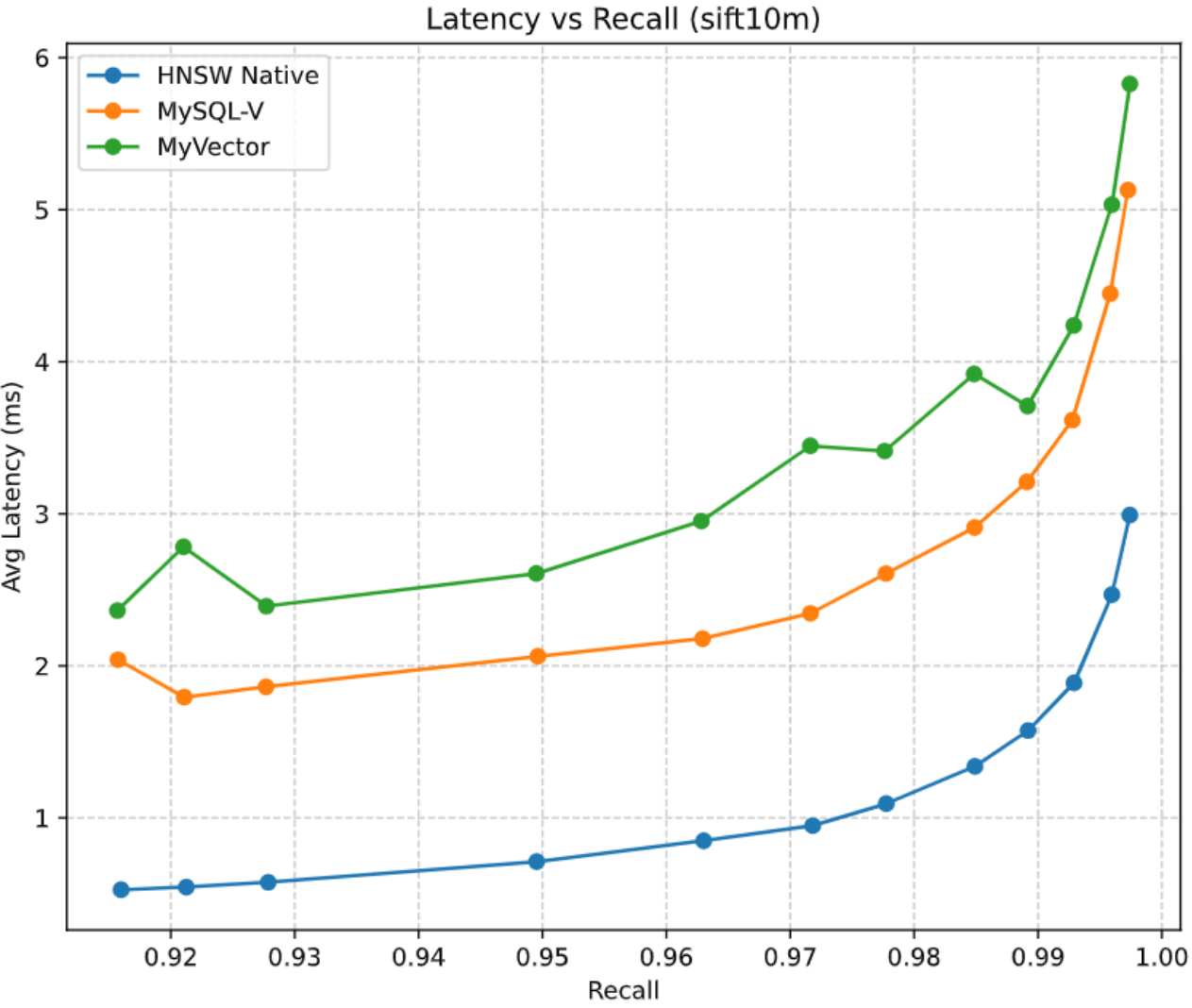

User Layer
User
Insert
Vector
Backend
Query
Vector
MySQL-V Memory Layer
MySQL
Traditional Indexing Modules
B+ Tree Index
Spatial Index
Full Text Index
Main Table
Vector Index
Auxiliary Table
Vector Index Interfaces
Search
Insert
Delete
Update
LSM structure
Mutable Segment
Immutable Segments
Flush
Background Thread
Merge
Background Thread
1
2
3
4
MySQL storage
LSM Persisted Data
Disk Storage Layer

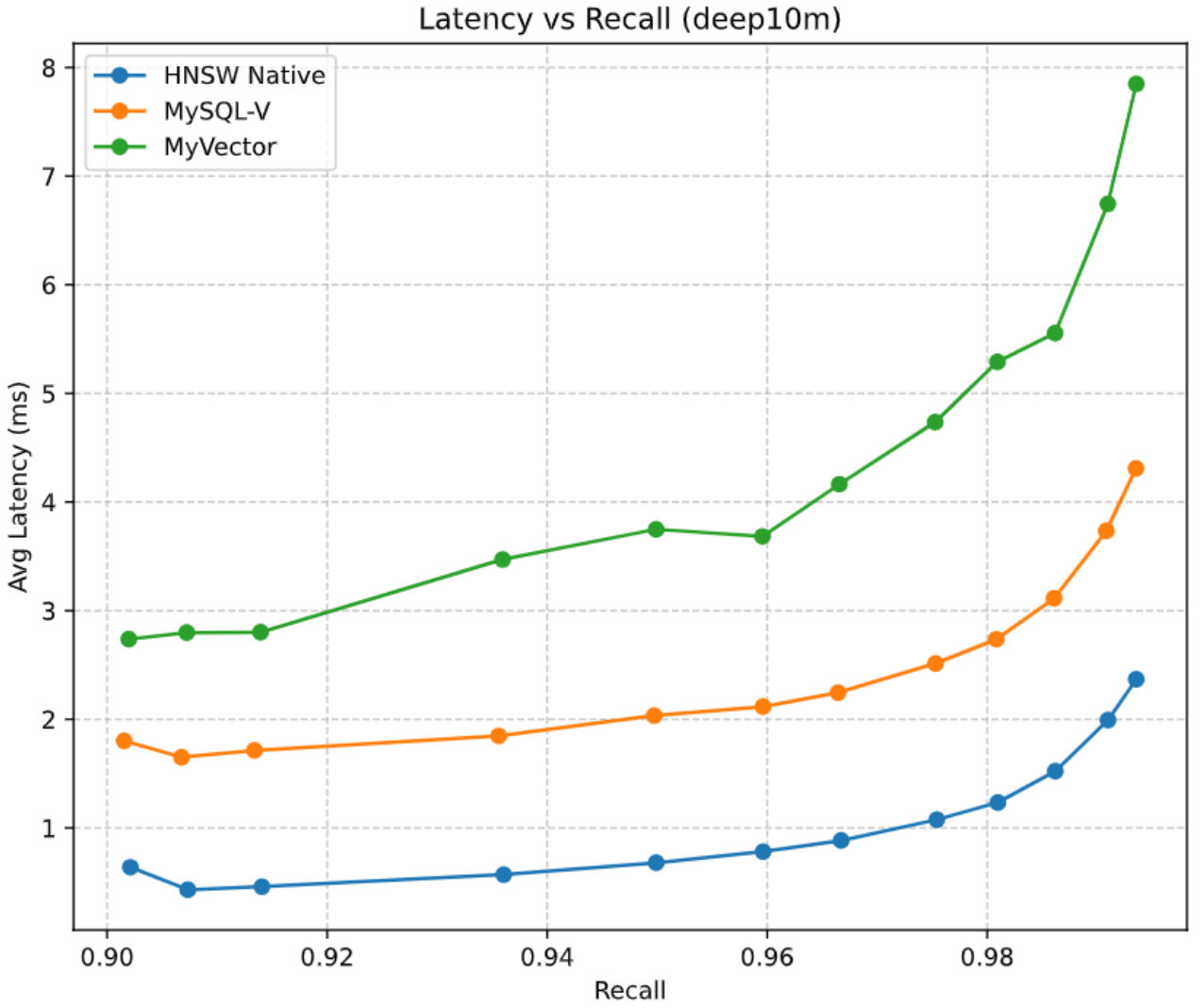
Intergrate vector search in MySQL





Intergrate vector search in MySQL





Thank you!