



Vector Databases

a deep dive

Guilherme Henriques, Alexandre Salamin, Bastien Schneider

Summary

01

Context

What the project is about

02

Concept

Vector databases explained

03

Comparison

How are vector databases different from normal DB

04

Project

Our project presented with a live demo

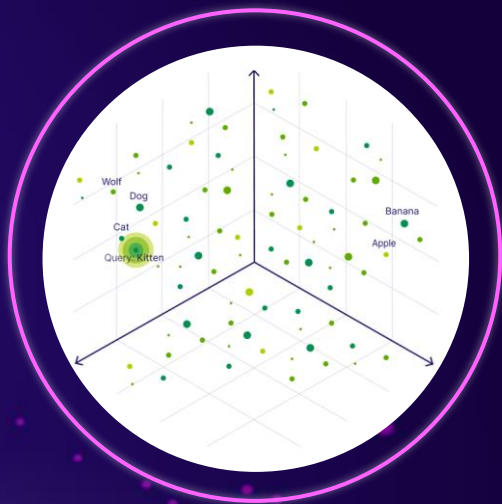


01

Context

Context

Vector DB



A way to approach databases
through machine learning

Recommendation



Our use case was a games
recommendation system



02

Concept

Concept

Vector Databases are composed of 2 elements :

Data

Row-collected information stored in any format (music, image, docx, csv, sql...)



Vectors

List of numbers created from the data through machine learning to mathematically represent the data on a graph



Concept



```
{  
  "firstName" : "Jean"  
  "name" : "Del"  
  "age" : 22  
}
```



```
{  
  23, 27,  
  10, 70  
}
```



Concept

Recommendation System

Finding closest elements based on other element



Search System

Finding closest elements based on text query



Image/Video Recognition

Finding closest elements based on img/vid query



All based on scores attributed by the vectors, a way to calculate the distance between them and a model to determine scores



03

Comparison

Comparison

	Vector Database	Traditionnal Database
Model	Store data in vector form	Structured data model (key-values, graph db, rows/columns)
Query	Similarity search algorithms (cosine) Fast regardless of size	Sql or other type Take longer if big database
UseCase	Recommendation, image and video search	Multiple (DW, ERP...)



04

Project

Project

Why Qdrant?

Self deployed

Can be built from scratch and deployed either on premise or through cloud



Editable code

Other vector databases are mostly premade (Pinecode for example) with no access to code

Free to use

Need to pay only for cloud hosting, otherwise completely free



Good documentation

Clear documentation with exemples.
Can quickly be ready to code

Project



Steam Dataset

Dataset with 80'000
videogames



Qdrant

Vector Database system



Python

Simpler programming
language for us



DEMO

What we actually did

The background is a deep purple color. It is decorated with numerous small, glowing purple dots scattered across the surface. In the top right and bottom left corners, there are intricate, glowing purple line patterns that resemble a wireframe or a complex network of fibers. The text is centered in the upper half of the image.

**Thank you
for your
attention!**