

# دوره آموزش Vector Database

[با رویکرد تصویر و متون فارسی]



طول دوره: ۶ الی ۸ ساعت

مدرسين: مسعود طاهري و حميدرضا حداد

## دوره آموزش Vector Database [با رویکرد تصویر و متون فارسی]: پایگاه داده برداری

(Vector Database) یک سیستم ذخیره سازی اطلاعات است که از بُردارهای عددی برای نمایش داده‌ها استفاده می‌کند. این بردارها معمولاً ویژگی‌های مختلفی از اطلاعات را نمایش می‌دهند. چند ویژگی مهم این نوع پایگاه داده عبارتند از:

۱. **قابلیت جستجو براساس بُردارها:** این پایگاه داده به کاربران این امکان را می‌دهد تا جستجوهای خود را براساس ویژگی‌های موجود در بردارها انجام دهند.

۲. **پشتیبانی از انواع مختلف داده:** این نوع پایگاه داده می‌تواند انواع مختلف داده‌ها را شامل تصاویر، صدا، فیلم و سایر اطلاعات پشتیبانی کند.

۳. **قابلیت مقایسه و تحلیل داده‌ها:** با استفاده از بردارها، این پایگاه داده امکان مقایسه و تحلیل داده‌ها را براساس ویژگی‌های آن‌ها فراهم می‌کند.

این ویژگی‌ها باعث می‌شود پایگاه داده برداری یک ابزار قدرتمند برای ذخیره، جستجو و دسترسی به اطلاعات مختلف باشد.

### مسعود طاهری: مدرس و مشاور ارشد SQL Server & BI

مدیر فنی پروژه‌های هوش تجاری (بیمه سامان، اوقاف، جین وست، هلدینگ ماهان و...) ، مدرس دوره‌های SQL Server و هوش تجاری در شرکت نیک‌آموز و نویسنده کتاب PolyBase در SQL Server



### حمیدرضا حداد: متخصص، مشاور و مدرس سیستم‌های هوش تجاری



## فصل اول: آشنایی با مبانی Vector Database

در طی این جلسه شما با مباحث زیر آشنا خواهید شد:

- بررسی انواع ساختارهای داده‌ای
- بررسی انواع بانک‌های اطلاعاتی
- Vector Database چیست؟
- Embedding Model اساس کار Vector Database ها
- Vector Database ها چگونه کار می‌کنند؟
- بررسی مرحله Indexing در Vector Database ها
- بررسی مرحله Querying در Vector Database ها

پردازش Unstructured Data همیشه یکی از چالش‌هایی بوده که خیلی از دوستانی که با این نوع داده‌های کار می‌کنند، با آن درگیر هستند.

برای مثال، من همیشه دوست داشتم که بتوانم برنامه‌ای بنویسم که امکان جستجو بر پایه عکس را داشته باشد. به طور خیلی ساده، تعدادی عکس در بانک اطلاعاتی خودم دارم، یک عکس جدید به‌عنوان ورودی به سیستم می‌دهم و می‌گویم عکس‌هایی شبیه به این را به من تحویل بده.

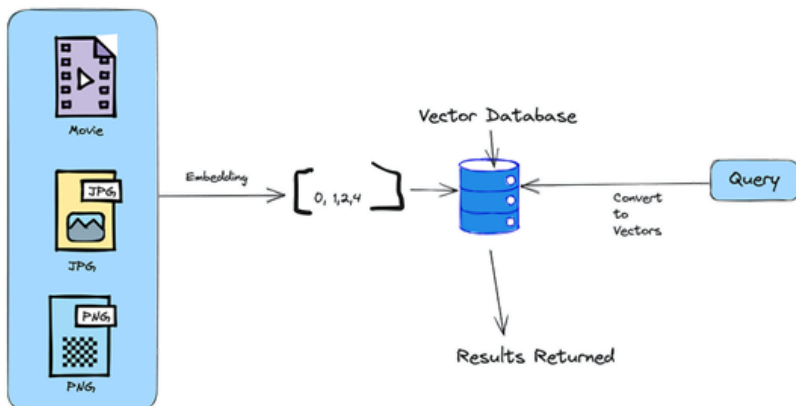
در گذشته برای انجام این کار باید درگیر هوش مصنوعی و... می‌شدیم، اما در حال حاضر انجام این کار با استفاده از Vector Database ها به راحتی آب خوردن است. برای انجام این کار کافی است:

### ذخیره سازی داده

- داده موردنظر (در اینجا عکس) را با استفاده از کتابخانه‌های Embedding به وکتور تبدیل نمایید.
- داده وکتور شده را در Vector Database ذخیره کنید تا ایندکس شود.

### جستجوی داده

- داده موردنظر (در اینجا عکس) را با استفاده از کتابخانه‌های Embedding به وکتور تبدیل نمایید.
- با استفاده از عملیات Similarity Search موجود در Vector Database ، جستجو داده وکتور را انجام داده و نزدیک‌ترین رکوردها به وکتور موردنظر را بازیابی نمایید.
- تریک می‌گویم! شما N تصاویر مشابه را استخراج کرده‌اید.
- تمام شد.



## فصل دوم: آشنایی با مبانی Vector Database

در طی این جلسه، شما با مباحث زیر آشنا خواهید شد:

- کاربردهای Vector Database ها
- LLM ها و Vector Database ها
- چند نمونه کاربرد واقعی از Vector Database ها
- معرفی چند Vector Database
- مزایا و معایب Vector Database ها
- بررسی Milvus به عنوان Vector Database
- بررسی Redis به عنوان Vector Database
- بررسی LanceDB به عنوان Vector Database

یکی از پایگاه داده‌های قوی‌ای که به عنوان Vector Database در سطح Enterprise فعالیت دارد، Milvus است. این پایگاه داده توانایی مدیریت و جستجوی داده‌های بُرداری را دارد. از نمونه کاربردهای Milvus می‌توان به امکان پیاده‌سازی جستجو براساس محتوای متن، تشخیص الگو، جستجو براساس محتوای تصویر، جستجو براساس محتوای صدا و... اشاره کرد.

از نمونه ویژگی‌های Milvus می‌توان به موارد زیر اشاره کرد:

- Open Source بودن آن
- پشتیبانی از انواع داده‌های بُرداری مانند وکتورها، ماتریس‌ها و تنسورها
- امکان جستجوی نزدیک‌ترین همسایه‌ها براساس فاصله یا شباهت کیفیتی
- امکان اضافه کردن، حذف کردن و به‌روزرسانی داده‌ها به صورت پویا
- پشتیبانی از معماری‌های مختلف مانند CPU، GPU
- امکان ایجاد ایندکس‌های بهینه برای بهبود عملکرد جستجوی داده‌ها
- و...



## فصل سوم: آشنایی با LanceDB

در طی این جلسه، ما به صورت عملی با بانک اطلاعاتی LanceDB کار خواهیم کرد. مباحث این جلسه شامل موارد زیر است:

- نحوه نصب و راه اندازی LanceDB به صورت Embedded Database
- آشنایی با نحوه ایجاد بانک اطلاعاتی در LanceDB
- آشنایی با نحوه ایجاد جدول در LanceDB
- آشنایی با نحوه انجام عملیات CRUD در LanceDB
- آشنایی با ایندکس Full Text در LanceDB
- کار با مثال‌های مناسب برای استفاده‌کنندگان از داشبورد

برای کار با LanceDB ما پکیج‌های مربوط به آن را در پایتون نصب و راه‌اندازی خواهیم کرد و با چند خط دستور، یک بانک اطلاعاتی در کنار Application خود ایجاد خواهیم کرد. به این نوع بانک‌های اطلاعاتی در اصطلاح Embedded Database می‌گویند.

LanceDB

Application

LanceDB

**Connect to a database**

Python   Typescript   Rust

```
import lancedb
import pandas as pd
import pyarrow as pa

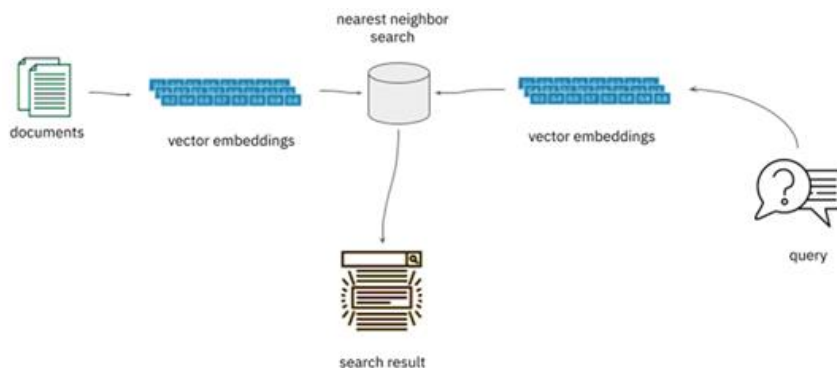
uri = "data/sample-lancedb"
db = lancedb.connect(uri)

# LanceDb offers both a synchronous and an asynchronous client.
# few operations that are only supported by the synchronous client
# (functions, full text search) but both APIs should soon be equivalent.

# In this guide we will give examples of both clients. In other
# typically only provide examples with one client or the other.
uri = "data/sample-lancedb"
async_db = await lancedb.connect_async(uri)
```

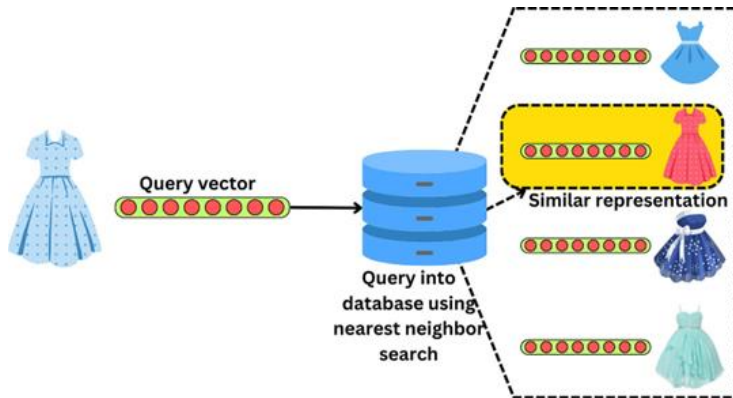
## فصل چهارم: پیاده‌سازی Semantic Search (جستجوی معنایی) در LanceDB

در طی این جلسه، ما به صورت عملی یک پیاده‌سازی یک مثال با استفاده از LanceDB برای جستجوی معنایی را انجام خواهیم داد. مباحثی که در این جلسه بررسی خواهد شد به شرح زیر است. (در حال ضبط)



## فصل پنجم: پیاده‌سازی Image Search (جستجوی تصویر)

در طی این جلسه، ما به صورت عملی یک پیاده‌سازی یک مثال با استفاده از Vector Database برای جستجوی تصویر را انجام خواهیم داد. مباحثی که در این جلسه بررسی خواهد شد به شرح زیر است. (در حال ضبط)





## نحوه مشاهده دوره چگونه است؟

این دوره آموزشی را می‌توانید در یک پلیر اختصاصی مشاهده فرمایید. به راحتی می‌توانید این نرم‌افزار را مناسب با سیستم عامل خود (ویندوز، مک، اندروید، لینوکس و یا وب) دانلود نصب کرده و پس از کپی کلید لایسنس داخل نرم‌افزار، محصول خریداری شده را تماشا کنید.

## صدور فاکتور رسمی چگونه است؟

در صورت تمایل به دریافت فاکتور رسمی، پیش از خرید خود با واحد فروش مجموعه (۰۲۱ - ۹۱ ۰۷ ۰۰ ۱۷) تماس حاصل نمایید. **شایان ذکر است، امکان صدور فاکتور رسمی پس از خرید آنلاین از سایت مجموعه به هیچ عنوان وجود نخواهد داشت.**



**آدرس:** تهران، یوسف آباد، میدان فرهنگ، خیابان ۳۳، پلاک ۲۹، زنگ ۲، دفتر نیک آموز  
**شماره تماس:** ۰۲۱ - ۹۱ ۰۷ ۰۰ ۱۷ | **موبایل فروش:** ۰۹۱۰ ۴۰۰۶ ۲۰۶