## ⌄ vector store in cx

```
from google.colab import userdata
hf_token=userdata.get('HUGGINGFACEHUB_ACCESS_TOKEN')
```

```
import time
```

```
start = time.time()
```

```
!pip install langchain chromadb openai tiktoken pypdf langchain_openai langchain-community
```

```
Requirement already satisfied: langchain in /usr/local/lib/python3.12/dist-packages (1.1.3)
Requirement already satisfied: chromadb in /usr/local/lib/python3.12/dist-packages (1.3.7)
Requirement already satisfied: openai in /usr/local/lib/python3.12/dist-packages (2.9.0)
Requirement already satisfied: tiktoken in /usr/local/lib/python3.12/dist-packages (0.12.0)
Requirement already satisfied: pypdf in /usr/local/lib/python3.12/dist-packages (6.4.2)
Requirement already satisfied: langchain_openai in /usr/local/lib/python3.12/dist-packages (1.1.4)
Requirement already satisfied: langchain-community in /usr/local/lib/python3.12/dist-packages (0.4.1)
Requirement already satisfied: langchain-core<2.0.0,>=1.1.2 in /usr/local/lib/python3.12/dist-packages (from langchain) (1.2.2
Requirement already satisfied: langgraph<1.1.0,>=1.0.2 in /usr/local/lib/python3.12/dist-packages (from langchain) (1.0.4)
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in /usr/local/lib/python3.12/dist-packages (from langchain) (2.12.3)
Requirement already satisfied: build>=1.0.3 in /usr/local/lib/python3.12/dist-packages (from chromadb) (1.3.0)
Requirement already satisfied: pybase64>=1.4.1 in /usr/local/lib/python3.12/dist-packages (from chromadb) (1.4.3)
Requirement already satisfied: uvicorn>=0.18.3 in /usr/local/lib/python3.12/dist-packages (from uvicorn[standard]>=0.18.3->chro
Requirement already satisfied: numpy>=1.22.5 in /usr/local/lib/python3.12/dist-packages (from chromadb) (2.0.2)
Requirement already satisfied: posthog<6.0.0,>=2.4.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (5.4.0)
Requirement already satisfied: typing-extensions>=4.5.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (4.15.0)
Requirement already satisfied: onnxruntime>=1.14.1 in /usr/local/lib/python3.12/dist-packages (from chromadb) (1.23.2)
Requirement already satisfied: opentelemetry-api>=1.2.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (1.39.1)
Requirement already satisfied: opentelemetry-exporter-otlp-proto-grpc>=1.2.0 in /usr/local/lib/python3.12/dist-packages (from
Requirement already satisfied: opentelemetry-sdk>=1.2.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (1.39.1)
Requirement already satisfied: tokenizers>=0.13.2 in /usr/local/lib/python3.12/dist-packages (from chromadb) (0.22.1)
Requirement already satisfied: pypika>=0.48.9 in /usr/local/lib/python3.12/dist-packages (from chromadb) (0.48.9)
Requirement already satisfied: tqdm>=4.65.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (4.67.1)
Requirement already satisfied: overrides>=7.3.1 in /usr/local/lib/python3.12/dist-packages (from chromadb) (7.7.0)
Requirement already satisfied: importlib-resources in /usr/local/lib/python3.12/dist-packages (from chromadb) (6.5.2)
Requirement already satisfied: grpcio>=1.58.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (1.76.0)
Requirement already satisfied: bcrypt>=4.0.1 in /usr/local/lib/python3.12/dist-packages (from chromadb) (5.0.0)
Requirement already satisfied: typer>=0.9.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (0.20.0)
Requirement already satisfied: kubernetes>=28.1.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (34.1.0)
Requirement already satisfied: tenacity>=8.2.3 in /usr/local/lib/python3.12/dist-packages (from chromadb) (9.1.2)
Requirement already satisfied: pyyaml>=6.0.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (6.0.3)
Requirement already satisfied: mmh3>=4.0.1 in /usr/local/lib/python3.12/dist-packages (from chromadb) (5.2.0)
Requirement already satisfied: orjson>=3.9.12 in /usr/local/lib/python3.12/dist-packages (from chromadb) (3.11.5)
Requirement already satisfied: httpx>=0.27.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (0.28.1)
Requirement already satisfied: rich>=10.11.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (13.9.4)
Requirement already satisfied: jsonschema>=4.19.0 in /usr/local/lib/python3.12/dist-packages (from chromadb) (4.25.1)
Requirement already satisfied: anyio<5,>=3.5.0 in /usr/local/lib/python3.12/dist-packages (from openai) (4.12.0)
Requirement already satisfied: distro<2,>=1.7.0 in /usr/local/lib/python3.12/dist-packages (from openai) (1.9.0)
Requirement already satisfied: jiter<1,>=0.10.0 in /usr/local/lib/python3.12/dist-packages (from openai) (0.12.0)
Requirement already satisfied: sniffio in /usr/local/lib/python3.12/dist-packages (from openai) (1.3.1)
Requirement already satisfied: regex>=2022.1.18 in /usr/local/lib/python3.12/dist-packages (from tiktoken) (2025.11.3)
Requirement already satisfied: requests>=2.26.0 in /usr/local/lib/python3.12/dist-packages (from tiktoken) (2.32.5)
Requirement already satisfied: langchain-classic<2.0.0,>=1.0.0 in /usr/local/lib/python3.12/dist-packages (from langchain-commu
Requirement already satisfied: SQLAlchemy<3.0.0,>=1.4.0 in /usr/local/lib/python3.12/dist-packages (from langchain-community)
Requirement already satisfied: aiohttp<4.0.0,>=3.8.3 in /usr/local/lib/python3.12/dist-packages (from langchain-community) (3.1
Requirement already satisfied: dataclasses-json<0.7.0,>=0.6.7 in /usr/local/lib/python3.12/dist-packages (from langchain-commun
Requirement already satisfied: pydantic-settings<3.0.0,>=2.10.1 in /usr/local/lib/python3.12/dist-packages (from langchain-comm
Requirement already satisfied: langsmith<1.0.0,>=0.1.125 in /usr/local/lib/python3.12/dist-packages (from langchain-community)
Requirement already satisfied: httpx-sse<1.0.0,>=0.4.0 in /usr/local/lib/python3.12/dist-packages (from langchain-community) (0
Requirement already satisfied: aiohappyeyeballs>=2.5.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp<4.0.0,>=3.8.3-
Requirement already satisfied: aiosignal>=1.4.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp<4.0.0,>=3.8.3->langcha
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp<4.0.0,>=3.8.3->langchain
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.12/dist-packages (from aiohttp<4.0.0,>=3.8.3->langch
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.12/dist-packages (from aiohttp<4.0.0,>=3.8.3->lang
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp<4.0.0,>=3.8.3->langcha
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.12/dist-packages (from aiohttp<4.0.0,>=3.8.3->langch
Requirement already satisfied: idna>=2.8 in /usr/local/lib/python3.12/dist-packages (from anyio<5,>=3.5.0->openai) (3.11)
```

```
!pip install huggingface_hub langchain-huggingface
```

```
Requirement already satisfied: huggingface_hub in /usr/local/lib/python3.12/dist-packages (0.36.0)
Collecting langchain-huggingface
  Downloading langchain_huggingface-1.2.0-py3-none-any.whl.metadata (2.8 kB)
```

```
Requirement already satisfied: filelock in /usr/local/lib/python3.12/dist-packages (from huggingface_hub) (3.20.0)
Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.12/dist-packages (from huggingface_hub) (2025.3.0)
Requirement already satisfied: packaging>=20.9 in /usr/local/lib/python3.12/dist-packages (from huggingface_hub) (25.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.12/dist-packages (from huggingface_hub) (6.0.3)
Requirement already satisfied: requests in /usr/local/lib/python3.12/dist-packages (from huggingface_hub) (2.32.5)
Requirement already satisfied: tqdm>=4.42.1 in /usr/local/lib/python3.12/dist-packages (from huggingface_hub) (4.67.1)
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.12/dist-packages (from huggingface_hub) (4.1
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.12/dist-packages (from huggingface_hub) (1.2.0)
Requirement already satisfied: langchain-core<2.0.0,>=1.2.0 in /usr/local/lib/python3.12/dist-packages (from langchain-huggingfa
Requirement already satisfied: tokenizers<1.0.0,>=0.19.1 in /usr/local/lib/python3.12/dist-packages (from langchain-huggingface)
Requirement already satisfied: jsonpatch<2.0.0,>=1.33.0 in /usr/local/lib/python3.12/dist-packages (from langchain-core<2.0.0,>=
Requirement already satisfied: langsmith<1.0.0,>=0.3.45 in /usr/local/lib/python3.12/dist-packages (from langchain-core<2.0.0,>=
Requirement already satisfied: pydantic<3.0.0,>=2.7.4 in /usr/local/lib/python3.12/dist-packages (from langchain-core<2.0.0,>=1.
Requirement already satisfied: tenacity!=8.4.0,<10.0.0,>=8.1.0 in /usr/local/lib/python3.12/dist-packages (from langchain-core<2
Requirement already satisfied: uuid-utils<1.0,>=0.12.0 in /usr/local/lib/python3.12/dist-packages (from langchain-core<2.0.0,>=1
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/python3.12/dist-packages (from requests->huggingface_h
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-packages (from requests->huggingface_hub) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/dist-packages (from requests->huggingface_hub) (2
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/dist-packages (from requests->huggingface_hub) (2
Requirement already satisfied: jsonpointer>=1.9 in /usr/local/lib/python3.12/dist-packages (from jsonpatch<2.0.0,>=1.33.0->langc
Requirement already satisfied: httpx<1,>=0.23.0 in /usr/local/lib/python3.12/dist-packages (from langsmith<1.0.0,>=0.3.45->langc
Requirement already satisfied: orjson>=3.9.14 in /usr/local/lib/python3.12/dist-packages (from langsmith<1.0.0,>=0.3.45->langcha
Requirement already satisfied: requests-toolbelt>=1.0.0 in /usr/local/lib/python3.12/dist-packages (from langsmith<1.0.0,>=0.3.4
Requirement already satisfied: zstandard>=0.23.0 in /usr/local/lib/python3.12/dist-packages (from langsmith<1.0.0,>=0.3.45->lang
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python3.12/dist-packages (from pydantic<3.0.0,>=2.7.4->l
Requirement already satisfied: pydantic-core==2.41.4 in /usr/local/lib/python3.12/dist-packages (from pydantic<3.0.0,>=2.7.4->la
Requirement already satisfied: typing-inspection>=0.4.2 in /usr/local/lib/python3.12/dist-packages (from pydantic<3.0.0,>=2.7.4-
Requirement already satisfied: anyio in /usr/local/lib/python3.12/dist-packages (from httpx<1,>=0.23.0->langsmith<1.0.0,>=0.3.45
Requirement already satisfied: httpcore==1.* in /usr/local/lib/python3.12/dist-packages (from httpx<1,>=0.23.0->langsmith<1.0.0,
Requirement already satisfied: h11>=0.16 in /usr/local/lib/python3.12/dist-packages (from httpcore==1.*->httpx<1,>=0.23.0->langs
Downloading langchain_huggingface-1.2.0-py3-none-any.whl (30 kB)
Installing collected packages: langchain-huggingface
Successfully installed langchain-huggingface-1.2.0
```

```python
from langchain_huggingface import HuggingFaceEndpointEmbeddings, HuggingFaceEndpoint
from langchain_community.vectorstores import Chroma
```

```python
from langchain_core.documents import Document
```

```python
# Create LangChain documents for IPL players

doc1 = Document(
        page_content="Virat Kohli is one of the most successful and consistent batsmen in IPL history. Known for his aggressive
        metadata={"team": "Royal Challengers Bangalore"}
    )
doc2 = Document(
        page_content="Rohit Sharma is the most successful captain in IPL history, leading Mumbai Indians to five titles. He's k
        metadata={"team": "Mumbai Indians"}
    )
doc3 = Document(
        page_content="MS Dhoni, famously known as Captain Cool, has led Chennai Super Kings to multiple IPL titles. His finishi
        metadata={"team": "Chennai Super Kings"}
    )
doc4 = Document(
        page_content="Jasprit Bumrah is considered one of the best fast bowlers in T20 cricket. Playing for Mumbai Indians, he
        metadata={"team": "Mumbai Indians"}
    )
doc5 = Document(
        page_content="Ravindra Jadeja is a dynamic all-rounder who contributes with both bat and ball. Representing Chennai Sup
        metadata={"team": "Chennai Super Kings"}
    )
```

```python
embeddings = HuggingFaceEndpointEmbeddings(
    repo_id="sentence-transformers/all-MiniLM-L6-v2",
    task="feature-extraction",
    huggingfacehub_api_token=hf_token,
)
```

```
vector_store=Chroma(
    embedding_function=embeddings,
    persist_directory="chroma_db",
    collection_name="players",
)
```

```
docs=[doc1,doc2,doc3,doc4,doc5]
```

```
# add documents
vector_store.add_documents(docs)
```

```
['0f31649f-7bcb-458a-89d1-c8611cf304bc',
 '5b516ede-f3d7-4555-bf1a-296ca26de267',
 'cfa12eb5-3e7d-485b-8452-de19a2d0677a',
 'e1b1a010-6011-4aa5-9714-4e5e5f333127',
 'bb1aa8e6-2009-4bd3-9fb3-8e242e4db6d9']
```

```
print(hf_token)
```

```
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
```

```
vector_store.get(include=['embeddings','documents','metadatas']) # to view the vector
```

```
{'ids': ['0f31649f-7bcb-458a-89d1-c8611cf304bc',
  '5b516ede-f3d7-4555-bf1a-296ca26de267',
  'cfa12eb5-3e7d-485b-8452-de19a2d0677a',
  'e1b1a010-6011-4aa5-9714-4e5e5f333127',
  'bb1aa8e6-2009-4bd3-9fb3-8e242e4db6d9'],
 'embeddings': array([[ 0.00994718,  0.06914335, -0.05147116, ..., -0.03543343,
          0.0128481 ,  0.01248287],
        [ 0.0012775 ,  0.03129857, -0.0237538 , ..., -0.0051836 ,
         -0.03280616,  0.02737708],
        [-0.10265914,  0.02650807,  0.02271501, ..., -0.03359747,
         -0.07984945, -0.01507713],
        [ 0.0212339 , -0.02468549, -0.0449436 , ..., -0.10995808,
          0.00572554,  0.09915374],
        [ 0.01873968,  0.04382843, -0.04304254, ..., -0.07801618,
         -0.07840687, -0.00304189]]),
 'documents': ['Virat Kohli is one of the most successful and consistent batsmen in IPL history. Known for his aggressive
batting style and fitness, he has led the Royal Challengers Bangalore in multiple seasons.',
  "Rohit Sharma is the most successful captain in IPL history, leading Mumbai Indians to five titles. He's known for his calm
demeanor and ability to play big innings under pressure.",
  'MS Dhoni, famously known as Captain Cool, has led Chennai Super Kings to multiple IPL titles. His finishing skills,
wicketkeeping, and leadership are legendary.',
  'Jasprit Bumrah is considered one of the best fast bowlers in T20 cricket. Playing for Mumbai Indians, he is known for his
yorkers and death-over expertise.',
  'Ravindra Jadeja is a dynamic all-rounder who contributes with both bat and ball. Representing Chennai Super Kings, his quick
fielding and match-winning performances make him a key player.'],
 'uris': None,
 'included': ['embeddings', 'documents', 'metadatas'],
 'data': None,
 'metadatas': [{'team': 'Royal Challengers Bangalore'},
  {'team': 'Mumbai Indians'},
  {'team': 'Chennai Super Kings'},
  {'team': 'Mumbai Indians'},
  {'team': 'Chennai Super Kings'}]}
```

```
# serach on similarity
result=vector_store.similarity_search(query="who is virat kohli",k=2)
```

```
print(result)
```

```
[Document(metadata={'team': 'Royal Challengers Bangalore'}, page_content='Virat Kohli is one of the most successful and consiste
```

```
# similarity with score
vector_store.similarity_search_with_score(query="who is virat kohli",k=2)
```

```
[(Document(metadata={'team': 'Royal Challengers Bangalore'}, page_content='Virat Kohli is one of the most successful and
consistent batsmen in IPL history. Known for his aggressive batting style and fitness, he has led the Royal Challengers
Bangalore in multiple seasons.'),
  0.4329704224586487),
 (Document(metadata={'team': 'Mumbai Indians'}, page_content="Rohit Sharma is the most successful captain in IPL history,
leading Mumbai Indians to five titles. He's known for his calm demeanor and ability to play big innings under pressure."),
  1.2250874042510986)]
```

```python
#meta-data filtering
vector_store.similarity_search_with_score(
    query="",
    filter={"team": "Chennai Super Kings"}
)
```

```
[(Document(metadata={'team': 'Chennai Super Kings'}, page_content='MS Dhoni, famously known as Captain Cool, has led Chennai
Super Kings to multiple IPL titles. His finishing skills, wicketkeeping, and leadership are legendary.'),
  1.8436005115509033),
 (Document(metadata={'team': 'Chennai Super Kings'}, page_content='Ravindra Jadeja is a dynamic all-rounder who contributes
with both bat and ball. Representing Chennai Super Kings, his quick fielding and match-winning performances make him a key
player.'),
  1.8909374475479126)]
```

```python
updated_doc1 = Document(
    page_content="Virat Kohli, the former captain of Royal Challengers Bangalore (RCB), is renowned for his aggressive leadersh
    metadata={"team": "Royal Challengers Bangalore"}
)
vector_store.update_document(document_id='0f31649f-7bcb-458a-89d1-c8611cf304bc', document=updated_doc1)
```

```python
# view document
vector_store.get(include=['embeddings','documents','metadatas'])
```

```
{'ids': ['0f31649f-7bcb-458a-89d1-c8611cf304bc',
  '5b516ede-f3d7-4555-bf1a-296ca26de267',
  'cfa12eb5-3e7d-485b-8452-de19a2d0677a',
  'e1b1a010-6011-4aa5-9714-4e5e5f333127',
  'bb1aa8e6-2009-4bd3-9fb3-8e242e4db6d9'],
 'embeddings': array([[-0.00233748,  0.05902077, -0.04774045, ..., -0.07264046,
          0.00276782, -0.00344092],
        [ 0.0012775 ,  0.03129857, -0.0237538 , ..., -0.0051836 ,
         -0.03280616,  0.02737708],
        [-0.10265914,  0.02650807,  0.02271501, ..., -0.03359747,
         -0.07984945, -0.01507713],
        [ 0.0212339 , -0.02468549, -0.0449436 , ..., -0.10995808,
          0.00572554,  0.09915374],
        [ 0.01873968,  0.04382843, -0.04304254, ..., -0.07801618,
         -0.07840687, -0.00304189]]),
 'documents': ["Virat Kohli, the former captain of Royal Challengers Bangalore (RCB), is renowned for his aggressive leadership
and consistent batting performances. He holds the record for the most runs in IPL history, including multiple centuries in a
single season. Despite RCB not winning an IPL title under his captaincy, Kohli's passion and fitness set a benchmark for the
league. His ability to chase targets and anchor innings has made him one of the most dependable players in T20 cricket.",
  "Rohit Sharma is the most successful captain in IPL history, leading Mumbai Indians to five titles. He's known for his calm
demeanor and ability to play big innings under pressure.",
  'MS Dhoni, famously known as Captain Cool, has led Chennai Super Kings to multiple IPL titles. His finishing skills,
wicketkeeping, and leadership are legendary.',
  'Jasprit Bumrah is considered one of the best fast bowlers in T20 cricket. Playing for Mumbai Indians, he is known for his
yorkers and death-over expertise.',
  'Ravindra Jadeja is a dynamic all-rounder who contributes with both bat and ball. Representing Chennai Super Kings, his quick
fielding and match-winning performances make him a key player.'],
 'uris': None,
 'included': ['embeddings', 'documents', 'metadatas'],
 'data': None,
 'metadatas': [{'team': 'Royal Challengers Bangalore'},
  {'team': 'Mumbai Indians'},
  {'team': 'Chennai Super Kings'},
  {'team': 'Mumbai Indians'},
  {'team': 'Chennai Super Kings'}]}
```

Start coding or generate with AI.