

ElememSDK Demo使用说明书-Python

1. 概述

本使用说明书详细介绍了如何使用Python语言编写的ElememSDK示例程序。该程序示范如何连接ElememSDK服务端，并执行索引的创建、训练、添加数据、查询数据、更新数据以及删除操作。

2. 环境准备

2.1 开发环境要求

- Python 3.7及以上版本
- ElememSDK Python客户端
- Docker（推荐使用Docker运行）

2.2 启动容器

```
# FAE提供的安装包，名称规格如下  
elemem-vector-engine.[2.0.8].tar.gz
```

```
# 解压  
tar xvf elemem-vector-engine.[2.0.8].tar.gz
```

```
# 解压后进入工程目录  
cd docker
```

```
# 目录中包含如下文件  
- docker-compose.yml
```

```
# 启动加速卡引擎  
sudo docker compose up -d server # -d是为了让容器在后台运行，不使用此参数会直接在当前运行，并直接打印日志到当前窗口。docker-compose.yml 中配置了本地端口8000映射到容器内端口8000  
可以通过sudo docker logs elemem_server查看容器启动的日志，容器内的服务是通过supervisor控制的。
```

关于compose使用的一些说明：

在旧版中，可能需要使用sudo docker-compose up -d server。在旧版docker时，docker-compose是一个独立的命令，属于Compose v1(2023年标记为deprecated)，新版docker(≥20.10.13)，compse v2(2020年推出的)可以作为一个插件安装，安装后compse是docker的一个子命令，建议使用最新版。

```
# 运行Demo镜像  
sudo docker compose run --rm client # --rm 代表退出后就删除本次创建的容器，请根据自己需
```

```
要修改运行参数
或者
sudo docker compose up -d client
sudo docker exec -it elemem_client /bin/bash
```

3. 程序运行

3.1 命令行参数说明

参数	描述	默认值
<code>--server</code>	ElememSDK服务器地址	localhost:7000
<code>--hdf5</code>	HDF5数据文件路径	必填
<code>--index</code>	索引名称	sift

3.2 运行示例

```
cd /root/hilbert/python
bash run.sh --server 127.0.0.1:7000 --hdf5 /root/hilbert/c++/SIFT_1M.hdf5 --
index sift
```

4. 功能说明

程序运行时，会执行以下一系列操作：

4.1 连接服务器

启动时自动连接指定ElememSDK服务端，连接成功会记录日志：

```
启动 Hilbert 客户端演示
服务器：<server>
HDF5 文件：<hdf5路径>
索引名称：<索引名>
```

4.2 索引操作

程序执行以下索引操作：

- 删除旧索引（如果存在）

- 创建新索引
- 查询所有索引

4.3 数据操作

执行索引的数据管理操作：

- 训练索引
- 添加向量
- 查询向量
- 执行搜索测试（计算召回率）
- 更新向量
- 随机查询搜索
- 删除向量
- 删除索引

每个步骤成功执行后会输出详细日志信息。

5. 常见问题与排查

- **连接失败：** 确认服务器IP及端口正确。
- **索引或数据操作失败：** 检查日志以确定具体错误原因。

6. 技术支持

如遇到任何问题或需要更多帮助，请联系ElememSDK技术支持团队。

本说明书旨在帮助您快速入门ElememSDK的Python示例程序。希望对您的开发工作有所助益！