# 自查询检索器原理与分类

#### 自查询检索器的概念

自查询检索器，顾名思义，是具有自我查询能力的检索器。具体来说，给定任何自然语言查询，检索器使用一个查询构造的大型语言模型（LLM）链来编写结构化查询，然后将该结构化查询应用于其底层的向量存储。这使得检索器不仅可以使用用户输入的查询与存储文档内容进行语义相似性比较，而且还可以从用户查询中提取关于存储文档元数据的过滤器，并执行这些过滤器。

#### 自查询检索器的分类

#### 1. Chroma自查询检索器

Chroma是一个用于构建嵌入式AI应用程序的数据库。在我们的示例中，我们将展示一个围绕Chroma向量存储的自查询检索器。

#### 2. MyScale自查询检索器

MyScale是一个集成向量数据库。你可以通过SQL和LangChain访问你的数据库。MyScale可以利用各种数据类型和函数进行过滤。无论你是扩大数据规模还是将系统扩展到更广泛的应用，MyScale都可以提升你的LLM应用的性能。

#### 3. Pinecone自查询检索器

在我们的演示中，我们将展示一个与Pinecone向量存储一起使用的自查询检索器。

#### 4. Qdrant自查询检索器

Qdrant是一个向量相似性搜索引擎。它提供了一个生产就绪的服务，具有便捷的API来存储、搜索和管理点 - 带有额外负载的向量。Qdrant针对扩展过滤支持进行了优化，使其更加实用。在我们的示例中，我们将展示一个围绕Qdrant向量存储的自查询检索器。

#### 5. Weaviate自查询检索器

创建Weaviate向量存储是首要步骤，我们希望为其添加一些数据。我们创建了一个包含电影摘要的小型演示文档集。

注意：自查询检索器需要你安装lark（使用pip install lark命令进行安装）。我们还需要weaviate-client包。

以上就是自查询检索器的基本概念和分类，希望能对你的学习和理解有所帮助。