数据的持久化设计

### 一、数据库选型

**数据库类型**：MySQL 8.x

**存储方式**：关系型表结构

**说明**：采用 MySQL 以支持结构化查询和事务控制，适合用户、评论、分析数据等持久化存储。

### 二、模块划分与持久化对象

|  |  |
| --- | --- |
| 模块 | 持久化对象 |
| 用户管理模块 | 用户信息（users） |
| 舆情收集模块 | 视频数据（videos）、评论数据（comments） |
| 舆情分析模块 | 情感分析结果（sentiments）、关键词提取（keywords） |
| 系统管理模块 | 通知、日志等（notifications、logs） |
| 用户管理模块 | 用户信息（users） |

### 三、数据表设计

1，用户表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 类字段名 | Python 类型 | 映射表字段名 | 数据库字段类型 | 说明 |
| id | int | id | INT PRIMARY KEY AUTO\_INCREMENT | 用户ID |
| username | str | username | VARCHAR(50) UNIQUE | 用户名 |
| password | str | password | VARCHAR(100) | 密码（加密） |
| role | str | role | VARCHAR(20) | 角色 |
| created\_at | datetime | created\_at | DATETIME | 注册时间 |

2.视频表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 类字段名 | 数据类型 | 映射表字段名 | 数据库字段类型 | 说明 |
| id | int | id | INT PRIMARY KEY AUTO\_INCREMENT | 视频ID |
| title | str | title | VARCHAR(200) | 标题 |
| description | str | description | TEXT | 描述 |
| upload\_time | datetime | upload\_time | DATETIME | 上传时间 |
| source | str | source | VARCHAR(50) | 平台来源 |

3. 评论表（comments）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 类字段名 | 数据类型 | 映射表字段名 | 数据库字段类型 | 说明 |
| id | int | id | INT PRIMARY KEY AUTO\_INCREMENT | 评论ID |
| video\_id | int | video\_id | INT | 视频ID |
| user | str | user | VARCHAR(100) | 评论者 |
| text | str | text | TEXT | 评论内容 |
| created\_at | datetime | created\_at | DATETIME | 评论时间 |

4. 情感分析表（sentiments）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 类字段名 | 数据类型 | 映射表字段名 | 数据库字段类型 | 说明 |
| id | int | id | INT PRIMARY KEY AUTO\_INCREMENT | 分析ID |
| comment\_id | int | comment\_id | INT | 评论ID |
| sentiment | str | sentiment | VARCHAR(20) | 情感分类 |
| confidence | float | confidence | FLOAT | 置信度 |
| analyzed\_at | datetime | analyzed\_at | DATETIME | 分析时间 |

5. 关键词表（keywords）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 类字段名 | 数据类型 | 映射表字段名 | 数据库字段类型 | 说明 |
| id | int | id | INT PRIMARY KEY AUTO\_INCREMENT | 关键词ID |
| video\_id | int | video\_id | INT | 视频ID |
| word | str | word | VARCHAR(50) | 关键词 |
| frequency | int | frequency | INT | 出现频率 |

6. 通知表（notifications）

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 类字段名 | 数据 类型 | 映射表字段名 | 数据库字段类型 | 说明 |
| id | int | id | INT PRIMARY KEY AUTO\_INCREMENT | 通知ID |
| user\_id | int | user\_id | INT | 用户ID |
| message | str | message | TEXT | 通知内容 |
| status | str | status | VARCHAR(20) | 状态 |
| created\_at | datetime | created\_at | DATETIME | 创建时间 |

### 四、实体关系图（ER 模型简述）

1.一个用户（users）可能接收多个通知（notifications)

2.一个视频（videos）有多个评论（comments）

3.每条评论（comments）对应一个情感分析记录（sentiments）

4.每个视频（videos）提取多个关键词（keywords）