

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）