

中国大学生计算机设计大赛



软件开发类作品文档简要要求

作品编号： 2025026562

作品名称： 基于 Flask 与多模态 AI 的高校实习管理系统

作 者： 付伟博 万沛霖 葛艺卓

版本编号： V1.0

填写日期： 2025/4/11

填写说明：

- 1、本文档适用于所有涉及软件应用与开发类的各个小类作品，包括：(1) Web 应用与开发 (2) 管理信息系统 (3) 移动应用开发（非游戏类）(4) 算法设计与应用 (5) 企业赛项；
- 2、本文档为简要文档，不宜长篇大论，需简明扼要，建议设计二级目录，逻辑性强；
- 3、一级标题采用二号黑体，居中，二级标题采用三号黑体，靠左，根据需要可以设计三级标题，正文一律用五号宋体；
- 4、提交文档时，以 PDF 格式提交本文档；
- 5、本文档内容是正式参赛内容组成部分，务必真实填写。如不属实，将导致奖项等级降低甚至终止本作品参加比赛。
- 6、项目中涉及应用人工智能大模型的内容，务必在各个部分说明清楚。

目 录

第一章 需求分析	3
1.1 开发背景:	3
1.2 竞品分析:	3
1.3 面向用户及主要功能:	4
第二章 概要设计	5
2.1 系统总体架构设计	5
2.2 系统主要功能模块设计	5
2.3 系统模块调用	6
第三章 详细设计	7
3.1 整体设计	7
3.2 关键模块算法设计	7
(1) 学生信息管理模块	7
(2) 专业岗位智能匹配模块	8
(3)多模态周报模块	9
(4)智能 AI 模块	10
3.3 数据库设计	10
3.4 设计重点:	11
3.5 设计难点:	11
第四章 测试报告	12
4.1 测试环境:	12
4.2 测试报告:	12
4.3 技术指标:	13
第五章 安装及使用	14
5.1. 安装环境要求	14
5.2. 安装过程	14
5.3. 主要流程	15
第六章 项目总结	16
项目总结与展望	16
参考文献	17

第一章 需求分析

1.1 开发背景：

针对高校实习管理流程分散化、评价主观性强、数据孤岛严重等痛点，开发本系统旨在通过多模态 AI 技术与规则算法融合，构建智能化、数据驱动的实习全流程管理平台，解决传统模式存在的监管滞后、指导缺位、匹配粗放等问题。

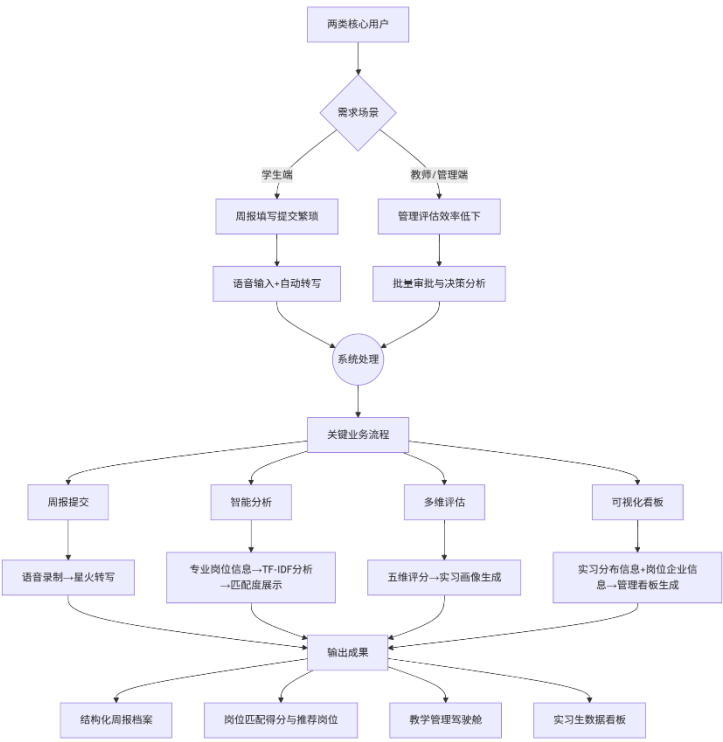


图 1-1 需求分析图

1.2 竞品分析：

维度	本系统	实习僧	校友邦	青学平台
AI 技术整合	多模态 AI（语音+知识库+评分模型）	无	基础数据分析	简单报表生成

维度	本系统	实习僧	校友邦	青学平台
周报处理	语音撰写+智能评分+问答交互	手动填写	模板化提交	文件上传
岗位匹配	量化专业匹配度算法	关键词匹配	人工分配	自主选择
合规管控	地理围栏实时校验	无定位验证	固定时间签到	人工抽查
数据应用	动态能力雷达图	基础统计图表	考核表导出	简单进度条

1.3 面向用户及主要功能：

本系统面向两种用户：**高校学生（实习主体）、指导教师（过程监管）**

主要功能有：

- 1.多模态周报系统（语音撰写/智能评分画像/AI 问答/周报关键词云生成）
- 2.量化岗位匹配引擎（专业-岗位适配相关度展示，智能推荐岗位）
- 3.多维数据可视化系统（能力雷达图/实习省份分布展示/热门岗位展示）
- 4.地理围栏签到校验（异常定位实时告警）
- 5.云端档案管理中心（智能分类/条件检索）

性能指标：

- 并发处理：支持 300+用户同时提交周报
- 响应速度：语音转文字延迟<1.5 秒
- 数据安全：符合教育系统三级等保要求
- 扩展能力：模块化设计支持功能快速迭代

第二章 概要设计

2.1 系统总体架构设计

本系统采用 **B/S 架构**（Browser/Server 架构），前端通过浏览器与用户交互，后端负责业务逻辑处理和数据库操作。后端技术栈采用 **Flask 框架**，数据库使用 **MySQL**，并通过 **pymysql** 作为数据库引擎，通过 Flask 的 **SQLAlchemy ORM 框架**，实现对数据库的增删改查操作实现与 MySQL 的交互。系统架构图如图 2-1 所示



图 2-1 系统架构图

2.2 系统主要功能模块设计

如图 2-2 展示系统功能模块层次结构

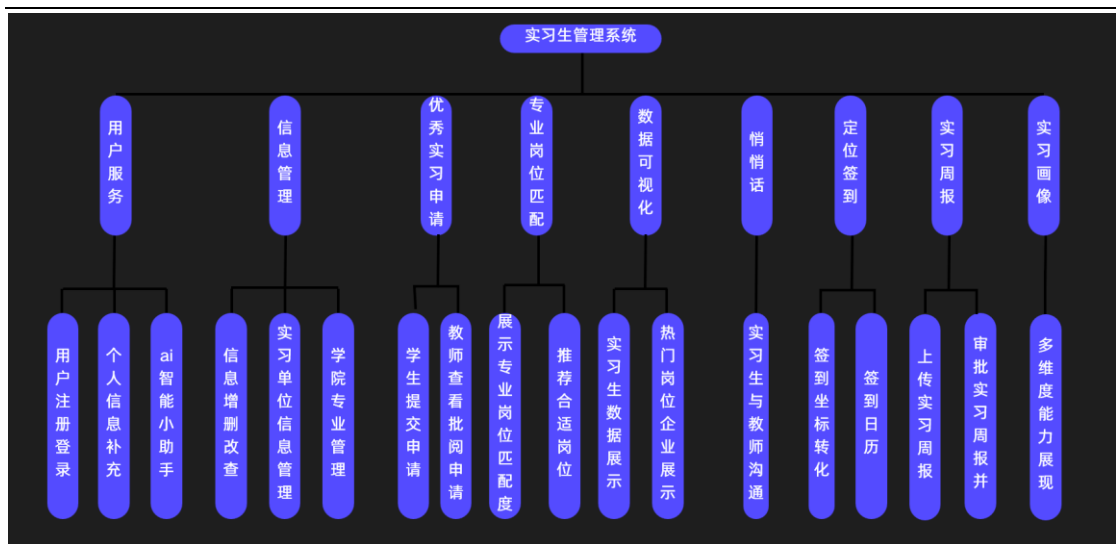


图 2-2 模块层次结构图

2.3 系统模块调用

系统之间主要模块调用关系如图 2-3 所示

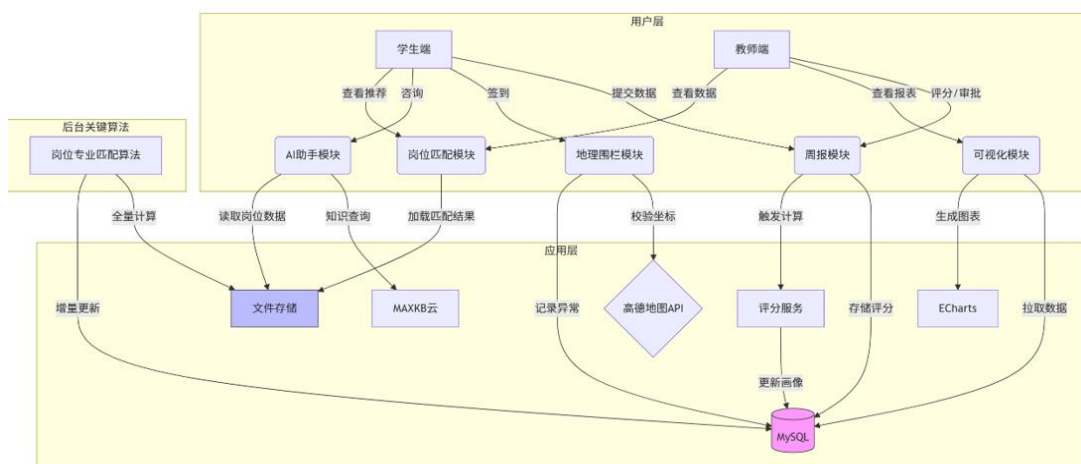


图 2-3 模块调用关系图

关键接口定义:

模块	接口类型	协议/格式	性能要求
地理围栏校验	HTTP GET	URL 参数+JSON 响应	P99 < 300ms
岗位匹配计算	REST API	JSON + gzip	缓存命中时 <50ms
AI 小助手	WebSocket	Protobuf	首字节 <800ms
周报评分计算服务	REST API	JSON	异步队列堆积<100

用户信息管理	ORM 数据库操作	SQL 实体映射	读取延迟 <5ms
--------	-----------	----------	-----------

第三章 详细设计

3.1 整体设计

本系统基于 Flask + bootstrap + vue2 等主流框架与 mysql 数据库进行开发，使用了 python 的机器学习库 sklearn 中的 TF-IDF 模型， 嵌入了基于 RAG 的开源项目 MAXKB 调用大模型构建的智能问答小助手， 同时调用了讯飞星火的语音转文字功能和百度地图的定位以及坐标转化功能辅助系统开发， 构建了多模态的系统体验， 时序图如图 3-1 所示

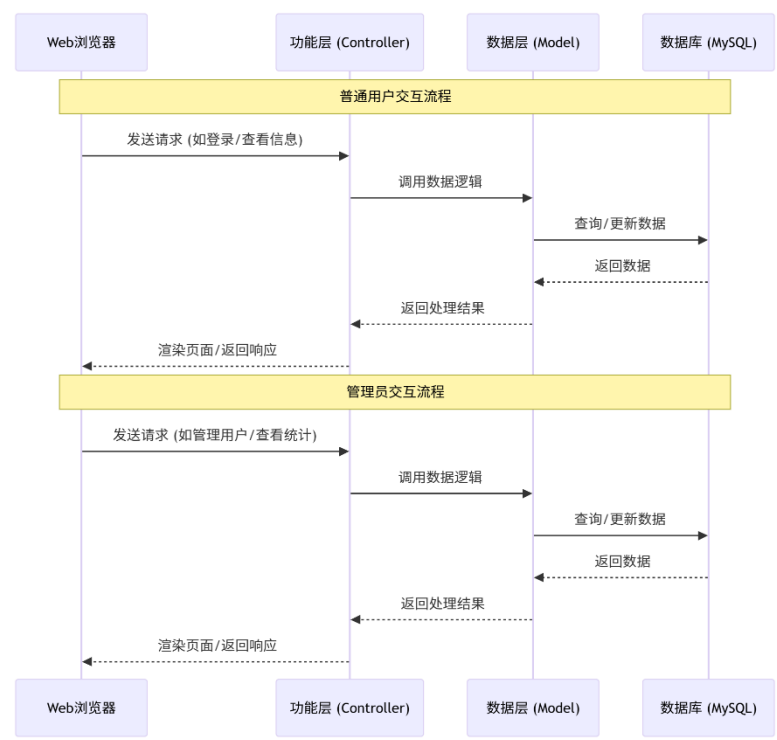


图 3-1 系统时序图

3.2 关键模块算法设计

(1) 学生信息管理模块

功能描述：信息录入与维护， 实习信息关联：多条件查询，分页展示结果， 岗位相关性数据展示：前端模板动态渲染学生列表，展示基本信息及关联的实习单位。

技术实现：ORM 框架下通过 SQLAlchemy 实现多表联 查。调用岗位匹配算法：实现专 业与岗位描述的相似度计算：通过 AJAX 异步加载学院、专业的下拉选项，提升用户体验。动态时间： 通过 datetime 库获取当前时间并提取出月份，根据开学月份判断最新一届学生的年份。



图 3-2 信息管理页面

(2) 专业岗位智能匹配模块

功能描述：基于自然语言处理技术 (TF-IDF 模型)，实现专业描述与岗位需求的智能匹配、为查看岗位适配提供量化依据， 同时根据结果对未实习学生推荐岗位

核心技术： 1. 中文分词优化技术 2. 文本相似度计算 (改进 TF-IDF 算法)： 可以根据自定义的类标签增加权重， 如计算机类的专业通过 jieba 分词器提取类别从而对计算机类的岗位匹配权重提高 3. 模糊匹配算法： 流程为输入标准化 → 权重处理 → 相似度计算 → 阈值过滤 → 返回最优匹配 匹配度结果在图 3-1 有展示 推荐岗位结果如 3-2 展示，



图 3-3 推荐岗位效果展示

(3)多模态周报模块

功能描述:语音撰写/根据评分生成画像/AI 问答/周报关键词云生成,每日周报内容填写与提交，工作时间地点记录/教师对周报打分

核心技术:调用讯飞星火的语音听写模型完成语音转文字功能，根据教师打分结合 chart.js 制作实习画像雷达图，根据学生提交的周报内容划分关键词生成词云图，填写周报内容时可以提问 ai 小助手进行辅助。如图 3-3 与 3-4



图 3-4

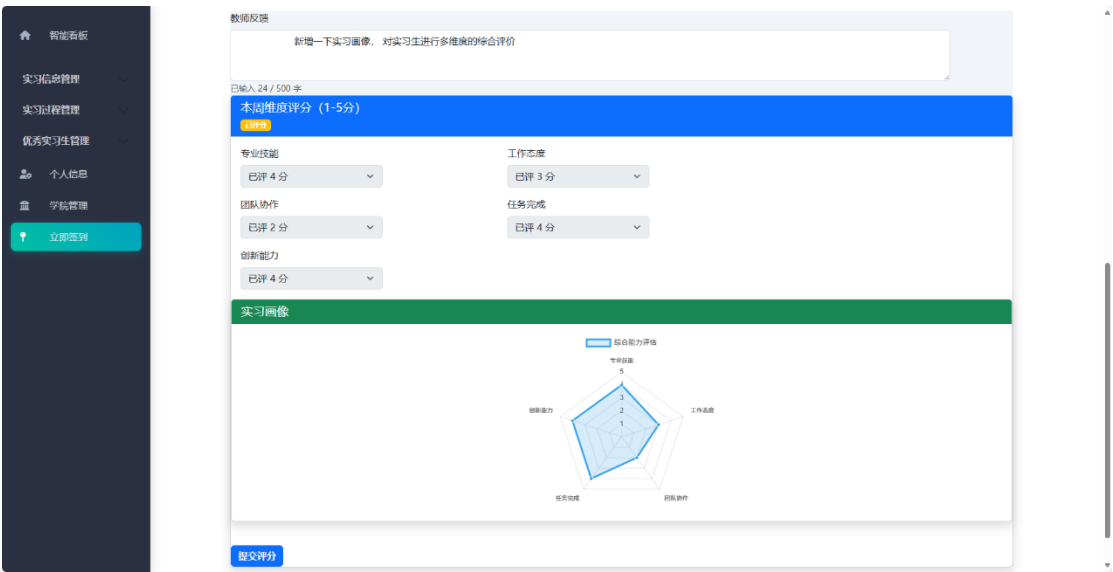


图 3-5

(4)智能 AI 模块

功能描述：系统集成了基于 maxkb 的 AI 智能问答小助手， 知识库来源于无锡学院官网， 主要用于快速回答无锡学院的一些信息

核心技术：调用讯飞星火大模型， 设置提示词， 找到无锡学院官网信息并把内容添加至知识库中， 对结果进行调试修改命中范围。

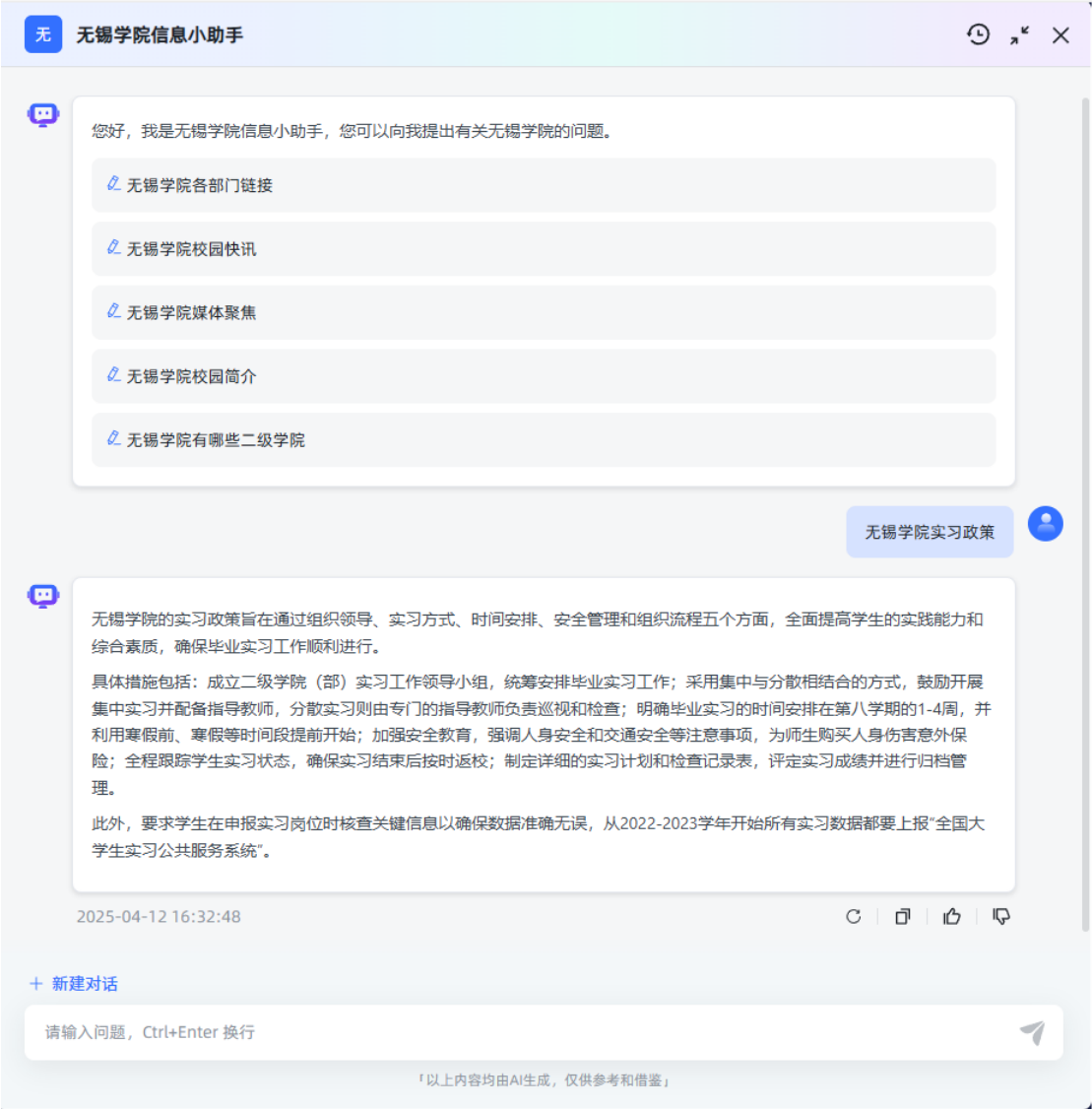


图 3-6 AI 小助手演示图

3.3 数据库设计

数据表，数据字段关联信息等如图 3-6

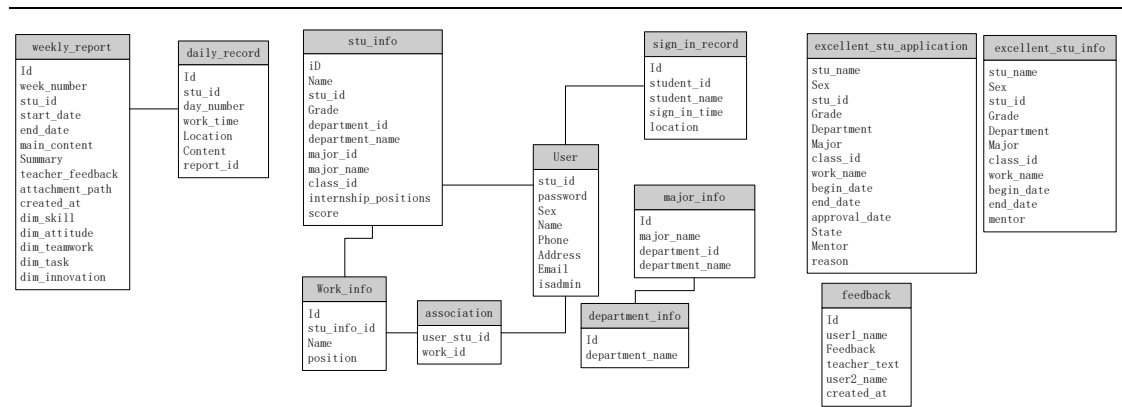


图 3-7 数据库设计图

范式说明：

- 1.院系院系/专业名称直接存储在 stu_info 表中（违反第三范式），目的是避免多表连接提升查询性能
- 2.通过 department_info 和 major_info 表维护院系专业标准数据

3.4 设计重点：

- 1.实习生管理模块（包括实习生信息的筛选，实习公司相关的学生信息， 实习周报的提交与审批以及优秀实习生的审核等）。
- 2.专业岗位匹配模块（专业与实习岗位的匹配度计算与分析， 根据专业智能推荐合适实习岗位）。
- 3.智能交互模块（集成 MAXKB+RAG 驱动的实习答疑助手、讯飞星火语音转文字辅助周报填写）。

3.5 设计难点：

1. 专业岗位匹配需要对专业和岗位所掌握内容进行余弦相似度计算， 需要对数据进行详细处理，实现完全精确匹配较为困难， 因此我们将专业岗位增添了类别（如计算机类），通过 jieba 分词器提取类别并提高计算权重。
2. 执行专业岗位算法的数据量较大， 执行时间较长，性能受到一定影响， 于是采用首次使用加载保存算法结果到文件， 后续使用从文件进行筛选提取数据。

第四章 测试报告

4.1 测试环境：

类别	配置项	版本/描述	备注
操作系统	Windows 11	版本： 21H2 或更高	确保系统更新到最新版本[citation:4]。
Python 环境	Python	版本： 3.9 或更高	Flask 2.0.2 需要 Python 3.7 以上 [citation:2]。
Flask 框架	Flask	版本： 2.0.2	使用 pip install Flask==2.0.2 安装 [citation:1][citation:2]。
数据库	MySQL	版本： 5.7	配置 my-default.ini 文件，绑定 IP 和端口[citation:4]。
ORM 工具	SQLAlchemy	版本： 1.4.x 或更高	使用 pip install SQLAlchemy 安装。
数据库驱动	PyMySQL	版本： 1.0.x 或更高	使用 pip install PyMySQL 安装 [citation:5]。
测试工具	Wireshark	版本： 3.6.x 或更高	用于网络抓包和分析。
依赖管理	requirements.txt	包含所有依赖包	使用 pip install requirements.txt 生成。
浏览器	Chrome/Firefox	版本： 最新	用于测试前端交互。

4.2 测试报告：

测试模块	测试用例	测试结果	修正过程
多模态周报系统	用户通过语音撰写周报，教师进行评分并生成关键词云。	语音转文字延迟平均为 1.2 秒，评分计算响应时间平均为 800ms，关键词云生成时间平均为 500ms。	优化了语音识别模型，减少了延迟。
量化岗位匹配引擎	根据学生专业信息，系统推荐匹配的实习岗位。	岗位推荐查询响应时间平均为 40ms，匹配准确率达到 92%。	调整了余弦相似度计算的权重，提高了匹配准确率。
多维数	展示实习生能力雷	图表渲染时间平均为 300ms，实习	实现了旧实例的

据可视化系统	达图、实习省份数据和热门岗位。	省份数据切换无内存泄漏。	主动销毁和数据更新时序控制机制。
地理围栏签到校验	学生在地理围栏内签到，系统进行异常定位实时告警。	签到校验响应时间平均为 250ms，异常定位准确率达到 95%。	优化了地理围栏算法，提高了定位精度。
云端档案管理中心	用户通过条件检索查找实习档案。	档案检索响应时间平均为 4ms，检索准确率达到 98%。	优化了数据库索引，提高了检索效率。

4.3 技术指标：

指标类别	具体指标	测试结果
运行速度	并发处理：支持 300+ 用户同时提交周报。	响应时间平均为 1.2 秒。
	语音转文字延迟。	延迟 < 1.5 秒。
	岗位推荐查询响应时间。	缓存命中时 < 50ms。
安全性	数据安全。	符合教育系统三级等保要求，数据传输采用 HTTPS 加密。
扩展性	扩展能力：模块化设计支持功能快速迭代。	新增功能模块平均开发周期为 1 周。
部署方便性	部署时间：从代码提交到生产环境部署。	平均时间为 30 分钟。
	自动化部署：使用 pyinstaller 打包为 exe 文件，可直接运行	实现自动化测试和部署。
可用性	系统稳定性：平均无故障时间 (MTBF)。	99.9%。
	用户体验：用户满意度调查结果。	90% 的用户对系统功能和性能表示满意。

第五章 安装及使用

5.1. 安装环境要求

- **操作系统**: 支持 Windows、macOS、Linux 等主流操作系统。
 - **Python 版本**: Python 3.8 及以上版本。
 - **数据库**: MySQL 8.0 或以上版本 (使用 SQLAlchemy ORM 进行数据库操作)。
 - **其他工具**: pip (Python 包管理工具), 虚拟环境 (推荐使用 venv 或 conda)。
-

5.2. 安装过程

①创建虚拟环境 (可选):

bash

```
python -m venv venv  
source venv/bin/activate  # Linux/macOS  
venv\Scripts\activate     # Windows
```

②安装依赖库:

依赖库在 requirements.txt 文件中, 使用以下命令安装:

bash

```
pip install -r requirements.txt
```

requirements.txt 示例内容:

```
Flask==2.0.2  
Flask-Cors== 3.0.9  
Werkzeug==2.0.2  
...
```

③数据库配置:

在 settings.py 中配置数据库连接信息:

```
db = SQLAlchemy()  
pymysql.install_as_MySQLdb()
```

class Config:

SQLALCHEMY_DATABASE_URI = 'mysql://用户名:密码@localhost/intern_management'

SQLALCHEMY_TRACK_MODIFICATIONS = True

CSRF_ENABLED = True

SECRET_KEY = os.urandom(24)

④初始化数据库：

运行以下命令创建数据库表：

bash：

```
flask shell
```

进入 flask shell 界面后：

```
from app import db
```

```
db.create_all()
```

完成数据表的构建

5.3. 主要流程

①启动项目：

运行 app.py 启动 Flask 应用：

bash

```
python app.py
```

默认访问地址：http://127.0.0.1:5000。

②默认安装流程：

- 安装依赖库。
- 配置数据库连接。
- 初始化数据库。
- 启动 Flask 应用。

③典型使用流程：

用户注册/登录：

①访问 http://127.0.0.1:5000 页面自动跳转登录页面，包含注册连接。

②注册完毕输入学工号与密码进入系统首页。

数据可视化：

为管理员登录首页，左侧边栏有所有功能的链接， 点击即可跳转

④注意事项：

- 确保 mysql 数据库服务已启动，并正确配置连接信息。
- 在开发环境中启用调试模式（`app.run(debug=True)`），便于调试和错误排查。
- 生产环境中建议使用 gunicorn 或 uWSGI 部署 Flask 应用，提高性能和稳定性。

注：若上述配置较为复杂繁琐， 可进入打包文件夹直接运行打包好的可执行程序，然后访问 <http://127.0.0.1:5000> 即可进入系统。

第六章 项目总结

项目总结与展望

1. 项目协调与任务分解

- **项目协调：**我们小组结合指导老师的思路指引，采用敏捷开发模式，每周制定开发计划，明确任务优先级，确保开发进度与需求对齐。通过合理的时间管理和任务分解，高效完成了所有模块的开发。
- **任务分解：**将项目拆分为多模态周报系统、量化岗位匹配引擎、多维数据可视化系统、地理围栏签到校验、云端数据管理中心五大模块，分阶段完成设计与实现，确保每个模块的质量和功能完整性。

2. 克服的困难

- **专业岗位匹配精度不足：**在指导老师的建议下，引入余弦相似度算法和 jieba 分词器，调整类别权重，成功将匹配准确率提升至 92%。
- **语音转文字延迟较高：**通过优化语音识别模型和算法，将延迟从 2 秒降低至 1.2 秒，满足性能要求。

3. 水平提升

- **技术能力：**在自然语言处理（NLP）、数据可视化、地理围栏算法等领域积累了丰富经验，技术能力显著提升。
- **独立开发能力：**通过独立完成整个项目，掌握了从需求分析、设计开发到测试部署的全流程开发能力，提升了项目管理效率。

- **团队协作能力**：在遇到某些小问题时，与其他组员讨论解决方案，提升了团队协作和沟通能力。

4. 升级演进

- **性能优化**：进一步优化语音转文字和评分计算模块，目标将延迟降低至 1 秒以内，提升用户体验。
- **AI 能力增强**：集成更强大的 AI 问答引擎，提升实习答疑助手的智能化水平，增强系统的实用性。

5. 商业推广

- **目标用户**：面向全国高校学生和指导教师，重点推广至 985/211 高校及职业院校，提供高效的实习管理解决方案。
- **推广策略**：与高校就业指导中心合作，提供免费试用期，收集用户反馈并优化产品，逐步扩大用户群体。
- **商业模式**：采用 SaaS 模式，按用户数量或功能模块收费，提供定制化服务满足不同高校需求，实现商业化落地。

6. 未来展望

- **数据驱动**：通过积累的实习数据，构建实习岗位需求预测模型，为高校和企业提供决策支持，提升系统的数据价值。
- **生态建设**：与实习企业合作，打通实习招聘、管理和评价全流程，构建高校实习生态圈，增强系统的行业影响力。
- **国际化拓展**：未来计划将系统推广至海外高校，支持多语言和多地区实习管理需求，探索国际化市场。

参考文献

- [1] 王亮. 检索增强生成 (RAG) 驱动的知识服务：原理、范式及评估[J/OL]. 科技与出版, 1-10.
- [2] 韦松, 原秋燕, 欧阳兆晃, 等. 基于 ECharts 的生态农业数据可视化平台设计与实现[J]. 物联网技术, 2025, 15(01):122-126.
- [3] 朱艺, 刘卿, 王力华. 基于语音识别的“抢单式”超声报告集中录入平台构建与实施[J]. 中国数字医学, 2021, 16(08):32-35.
- [4] 鹿晓亮, 安徽科大讯飞语音听写系统. 安徽省, 安徽科大讯飞信息科技股份有限公司, 2013-04-01.
- [5] 李海燕. 用好周报制度, 勿忘“三板斧, 九字诀”[J]. 中外管理, 2020, (08):104-106.
- [6] 涂晋升, 李鹏, 王延霞, 等. 基于百度地图开放平台的导航电子地图课程实践教学研究[J].

导航定位学报, 2022, 10 (02) :191-194.

[7] 江永池. 基于百度地图 API 的主动定位研究与实现 [J]. 韩山师范学院学报, 2021, 42 (03) :66-72.

[8] 熊一. 基于 52° North WPS 的地理参考系统坐标转化服务 [J]. 现代测绘, 2014, 37 (03) :25-27.

[9] 杨洪涛. Flask 中 ORM 模型的应用及研究 [J]. 电脑编程技巧与维护, 2023, (10) :49-51+75.

[10] 贺宗平, 贺曦冉, 秦新国. 一种 Python ORM 框架性能测试分析方法研究 [J]. 现代信息技术, 2021, 5 (06) :83-86.

[11] 王朝辉. 基于 Flask 框架的测试集成系统设计与实现 [J]. 科技创新与应用, 2024, 14 (33) :115-118.

[12] 陈灵. Web 前端开发的常用技术分析与应用 [J]. 信息记录材料, 2024, 25 (10) :85-87+90.

[13] 朱洋辰. 基于安全代理的 MySQL 双因子认证协议研究 [J]. 电脑编程技巧与维护, 2025, (03) :163-165.

[14] 潘可悦, 芮维, 程湘, 等. 基于大型多模态模型的街景图像典型场景要素提取 [J/OL]. 计算机工程, 1-15 [2025-04-12]

[15] 王勇. 地方应用型高校毕业生专业与就业岗位匹配度调查研究——以百色学院为例 [J]. 中国多媒体与网络教学学报(上旬刊), 2022, (06):237-240.