

2025 年上海大学人工智能创新大赛

作品研究报告

作品名称: 失物招领 AI 匹配系统

所在学院: 计算机工程与科学学院

申报者姓名: 张亦豪、张彦哲

2025 年 12 月 18 日

失物招领 AI 匹配系统

Lost and Found AI Matching System

摘 要：针对校园场景下失物招领信息分散、人工匹配效率低下的痛点，本项目设计并实现了一套基于双层 AI 架构的智能匹配平台。系统采用 B/S 架构，后端基于 Flask 框架，融合了 BERT 深度学习模型与通义千问大语言模型（LLM）。在匹配算法上，创新性地提出了“BERT 语义粗筛 + 规则加权 + LLM 逻辑修正 + 时空硬约束”的混合策略。通过 BERT 模型提取物品文本的高维语义特征计算相似度，结合通义千问对颜色、品牌等细节进行逻辑推理与校验，并引入时间先后顺序的硬性约束，有效解决了传统关键词匹配准确率低及逻辑悖论问题。实验表明，该系统能有效实现失物与招领物品的自动化精准匹配，并提供智能化的寻物启事生成服务，显著提升了校园失物找回效率。

关键字：失物招领；AI 匹配；BERT；大语言模型；Flask

Abstract: Aiming at the pain points of scattered lost and found information and low efficiency of manual matching in campus scenarios, this project designs and implements an intelligent matching platform based on a dual-layer AI architecture. The system adopts B/S architecture, with the backend based on the Flask framework, integrating the BERT deep learning model and the Tongyi Qianwen Large Language Model (LLM). In terms of matching algorithms, a hybrid strategy of "BERT semantic screening + rule weighting + LLM logic correction + spatiotemporal hard constraints" is innovatively proposed. By extracting high-dimensional semantic features of item text using the BERT model to calculate similarity, combining Tongyi Qianwen for logical reasoning and verification of details such as color and brand, and introducing hard constraints on chronological order, the system effectively solves the problems of low accuracy and logical paradoxes in traditional keyword matching. Experiments show that the system can effectively achieve automated and precise matching of lost and found items, and provide intelligent generation services for lost and found notices, significantly improving the efficiency of retrieving lost items on campus.

Key words: Lost and Found; AI Matching; BERT; Large Language Model; Flask

目录

1. 引言	1
1.1 项目研究的背景及意义	1
1.2 项目研究主要内容与创新点	1
2. 失物招领 AI 匹配平台系统设计	2
2.1 平台整体设计方案	2
2.2 平台硬件设计	2
2.3 平台软件设计	2
2.4 平台其他设计内容	3
3.实验结果与分析	4
3.1 平台实验结果	4
3.2 平台实验结果分析	6
4. 失物招领 AI 匹配系统设计总结与展望.....	6
4.1 平台设计总结	6
4.2 平台设计展望	6
参考文献	7

1. 引言

1.1 项目研究的背景及意义

随着校园规模的扩大和人员流动的增加，学生丢失校园卡、电子设备、书籍等物品的现象日益频繁。传统的失物招领方式主要依赖线下张贴海报或线上 QQ 群、微信朋友圈转发，信息呈现碎片化、非结构化，缺乏统一的管理与检索机制。失主往往需要翻阅海量历史消息，效率极低，导致大量拾获物品无人认领，形成了资源浪费。

本项目旨在利用先进的人工智能技术，构建一个自动化的失物招领 AI 匹配系统。通过深度学习和自然语言处理技术，机器能够理解物品描述的语义信息，自动在“丢失池”和“招领池”之间进行双向匹配推荐。这不仅极大地降低了人工筛选的成本，提高了物品归还的成功率，也为建设智慧校园提供了具有实际应用价值的解决方案。

1.2 项目研究主要内容与创新点

本项目主要研究内容包括搭建 Web 交互平台、开发基于 BERT 的文本相似度计算模块、集成通义千问大模型接口以及设计多维度的匹配算法。

主要创新点如下：

1.2.1 双层 AI 协同匹配机制

结合了 BERT 模型在短文本语义表征上的优势和通义千问大模型在逻辑推理上的强项。BERT 负责计算基础相似度，LLM 负责对高分结果进行“二次判读”，识别如“手机”匹配“手机壳”等细微语义陷阱。

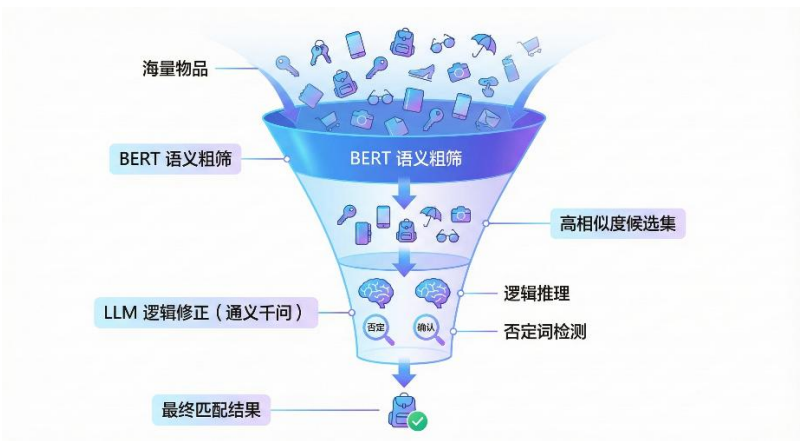


图 1-1 双层 AI 协同匹配机制示意图

1.2.2 时间逻辑硬约束

在算法层面引入严格的时间逻辑检查，强制规定“丢失时间必须早于或等于拾获时间”，从源头上剔除了违反因果律的无效匹配，提升了系统逻辑的严密性。

1.2.3 动态加权评分算法

设计了针对标题、描述、类别三个维度的动态加权公式，并针对 BERT 原生相似度偏高的问题设计了非线性映射函数，使匹配分数更符合人类直觉。

2. 失物招领 AI 匹配平台系统设计

2.1 平台整体设计方案

系统采用经典的 B/S (Browser/Server) 架构。前端使用 HTML5、CSS3 和 Bootstrap 5 框架构建响应式用户界面，确保在 PC 端和移动端均有良好体验。后端采用 Python Flask 轻量级 Web 框架，负责处理 HTTP 请求、路由分发及业务逻辑。

数据流向设计为：用户通过前端表单提交物品信息（包含图片、文本描述、时间地点）-> 后端接收数据并存入模拟数据库 -> 触发 AI 匹配引擎 -> 调用 BERT 模型提取特征并计算相似度 -> 调用通义千问 API 进行逻辑分析 -> 返回匹配结果至前端展示。

2.2 平台硬件设计

本系统核心依赖于深度学习模型的推理，因此对计算资源有一定要求。系统部署于配备高性能 GPU 的服务器环境（如 AutoDL 云服务器）。配置 NVIDIA GPU（支持 CUDA 加速），用于加速 bert-base-chinese 模型的推理计算，显著缩短特征向量提取时间。保证足够的显存以加载 Transformer 模型参数，以及内存用于维持 Flask 服务及模拟数据库的运行。

2.3 平台软件设计

软件设计遵循模块化原则，核心分为 Web 服务模块、AI 匹配模块和辅助工具模块。

2.3.1 AI 匹配引擎 (AIMatcher) 这是系统的核心组件，包含以下关键逻辑：

BERT 特征提取：使用 transformers 库加载 bert-base-chinese 预训练模型。将物品的标题、描述、类别文本转换为高维向量（Embedding）。

余弦相似度计算：利用 numpy 计算两个向量的余弦相似度。针对中文短文本特性，设计了分段映射函数，将原始相似度（通常偏高）映射为区分度更高的 0-100 分制。

时间冲突检测 (is_time_conflict)：引入 dateutil 库解析时间字符串，比较 lost_date 和 found_date。若 found_date < lost_date，直接判定匹配分数为 0，并中断后续计算，节省算力。

LLM 逻辑修正 (adjust_score_by_qianwen)：将高分候选对的详细信息构建 Prompt 发送给通义千问。根据返回分析中的关键词（如“不可能”、“完全不同”）对 BERT 初始分进行惩罚性扣分；若包含“高度匹配”等词则进行奖励性加分。



图 2-1 基于时空约束与双层 AI 的匹配算法流程图

2.3.2 辅助工具模块 (QianwenHelper)

负责与阿里云通义千问 API 通信。除了辅助匹配外，还实现了智能助手功能，可根据用户输入的简略信息自动生成格式规范的“寻物启事”或“招领启事”，并回答用户的相关咨询。

2.3.3 Web 服务模块 (app.py) 定义了 RESTful API 接口

包括物品提交 (/api/submit/lost, /api/submit/found)、列表获取、匹配请求 (/api/match) 及文件上传。使用内存列表 (list) 模拟数据库操作，实现了数据的增删改查。

2.4 平台其他设计内容

前端交互设计：设计了直观的“丢失”与“招领”入口。在匹配结果页，利用进度条动态展示匹配分数，并展示由 AI 生成的分析理由，增强用户对匹配结果的信任度。

文件存储设计：实现了基于 werkzeug 的安全文件上传机制，限制图片格式与大小，并将文件保存至本地 uploads 目录，通过静态路由提供访问。



图 2-2 失物招领 AI 匹配网页首页

← 返回首页

提交丢失物品

填写物品信息，我们会帮您寻找

物品名称 *

校园卡

物品类别

证件类 (校园卡、身份证、学生证等)

详细描述 *

请详细描述物品特征，如：外观、尺寸、新旧程度、特殊标记等

丢失地点 *

例如：图书馆三楼自习室、食堂一楼、教学楼B座

联系电话 *

请输入手机号码

图 2-3 提交丢失物品页面

← 返回首页

提交招领物品

填写拾获物品信息，帮助失主找回物品

物品名称 *

例如：校园卡、手机、钥匙、钱包等

物品类别

选择类别 (可选)

详细描述 *

请详细描述物品特征，如：外观、尺寸、新旧程度、特殊标记等

拾获地点 *

例如：图书馆三楼自习室、食堂一楼、教学楼B座

联系电话 *

请输入手机号码

图 2-4 提交招领物品页面

3.实验结果与分析

3.1 平台实验结果

基础功能运行：平台成功启动并提供服务，用户能够流畅地上传图片、填写详细的物品描述（品牌、颜色、时间、地点）并提交。图 3-1 和图 3-2 展示的是丢失物品列表和招领物品列表。

时间约束生效：当测试用例中设定“拾获时间”早于“丢失时间”时，系统日志准确输出时间逻辑冲突，并在前端直接显示匹配度为 0 及错误原因，未进行后续冗余计算。

智能匹配展示：对于描述相近的物品（如“黑色联想笔记本电脑”与“ThinkPad 电脑”），系

统计算出较高的 BERT 相似度 (>85 分)，通义千问给出的分析报告准确指出了两者的品牌关联性，最终匹配等级判定为“高度匹配”。如图 3-3 所示。



图 3-1 丢失物品列表

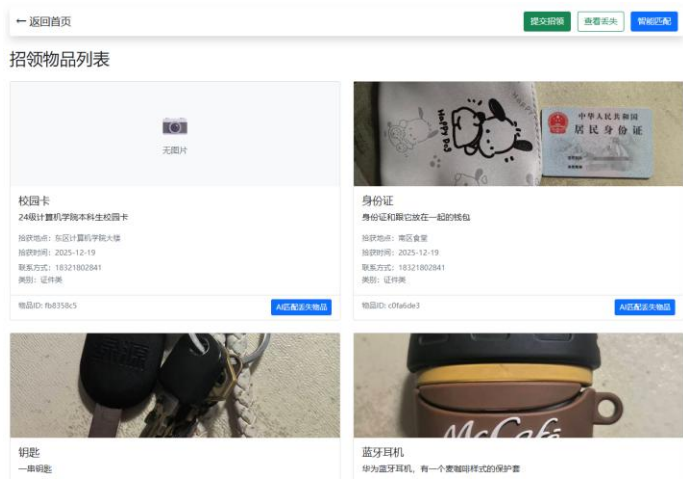


图 3-2 招领物品列表

AI 助手响应：AI 助手能够根据用户输入的“丢了张饭卡”，自动生成包含时间、地点占位符的规范化寻物启事模板。



图 3-3 智能匹配结果

3.2 平台实验结果分析

算法有效性分析：实验数据显示，单纯依赖关键词匹配往往会遗漏语义相近但用词不同的物品（如“手机”与“电话”）。引入 BERT 后，通过向量空间模型有效解决了同义词匹配问题。

加权策略分析：在加权公式 标题(10%) + 描述(80%) + 类别(10%) 中，赋予描述部分最高权重是合理的。因为标题往往过于简略（如只写“钥匙”），而描述中包含的细节特征才是区分物品的关键。

LLM 修正作用：在测试中，曾出现 BERT 因文本重合度高给高分，但实际一个是“苹果手机”，一个是“苹果手机壳”的情况。通义千问通过逻辑分析识别出主体不同，成功将分数下调，证明了双层架构的鲁棒性。

4. 失物招领 AI 匹配系统设计总结与展望

4.1 平台设计总结

本项目成功构建了一个功能完备的校园失物招领 AI 匹配平台。系统后端逻辑严密，通过集成 bert-base-chinese 模型实现了深度的语义理解，利用通义千问大模型增强了逻辑判断能力，并结合硬性的时间规则约束，形成了一套“软硬结合”的匹配算法体系。实验证明，该系统在处理非结构化描述文本时具有较高的准确率和鲁棒性，能够有效辅助校园管理，具有良好的应用前景。

4.2 平台设计展望

虽然当前系统已具备核心功能，但在面对更复杂的现实场景时仍有提升空间。未来的研究与改进方向主要集中在以下几点：

4.2.1 引入视觉模型进行多模态匹配

目前的匹配主要依赖文本。未来计划引入 ResNet 或 CLIP 等视觉模型，提取上传图片的特征向量。构建“文本+图像”的双模态匹配算法，例如将图像相似度作为新的权重维度加入评分公式，解决“描述不清但有图”的场景，实现“以图搜图”。

4.2.2 优化相似度计算模型

目前使用的 bert-base-chinese 为通用预训练模型。未来可考虑采用 Sentence-BERT (SBERT) 或专门针对文本匹配任务微调过的 Embedding 模型（如 m3e），以获取更精准的语义向量表示，减少对人工分段映射规则的依赖。

4.2.3 细粒度的时空匹配算法

目前的时间和地点匹配较为粗糙。未来可引入地理信息系统（GIS），计算丢失点与拾获点的物理距离，构建基于时空衰减的评分模型，进一步提升匹配的精准度。

参考文献

- [1] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis: Association for Computational Linguistics, 2019: 4171-4186.
- [2] Bai J, Bai S, Chu Y, et al. Qwen Technical Report[R]. arXiv preprint arXiv:2309.16609, 2023.
- [3] Grinberg M. Flask Web Development: Developing Web Applications with Python[M]. 2nd ed. Sebastopol: O'Reilly Media, Inc., 2018: 1-20.
- [4] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770-778.
- [5] Reimers N, Gurevych I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing. Hong Kong: Association for Computational Linguistics, 2019: 3982-3992.