

基于智能交互的个性化语音问答学习辅导系统设计

贾一帆¹, 王赞普¹, 孙嘉美¹, 贾涵淼¹, 韩奕博¹, 郝振²

¹河南工业大学人工智能与大数据学院, 河南郑州, 中国

²科大讯飞股份有限公司, 安徽合肥, 中国

【摘要】随着人工智能技术的进步, 其已广泛应用于教育行业, 深刻重塑了语音问答教育的服务模式与效率边界。但是现有的学习机以触控为主, 语音功能薄弱, 难以响应实时疑问, 缺乏自然对话感; 在线教育平台侧重录播或直播, 实时答疑能力有限, 无法针对个体即时卡点突破; 普通智能音箱的教育功能停留在浅度问答, 缺乏学科深度与体系化学习闭环。为了解决这些问题, 提出了一种基于智能交互的个性化语音问答学习辅导系统。通过整合前沿的自然语言处理与语音交互技术, 将人工智能的智能化服务特性融入家庭教育场景, 成为人工智能技术赋能教育领域的典型实践范例。为个性化学习与高效辅导提供了新的实现路径, 有效突破了传统教育模式在时空限制与资源分配不均方面的瓶颈。

【关键词】智能交互; 语音问答; 个性化设计; 学习辅导; 自然语言处理

1. 引言

近年来, 人工智能技术发展迅速, 其影响力已遍及全球多个领域, 技术革新与应用落地的速率持续刷新着人们对智能系统的认知边界[1]。人工智能技术的进步正驱动各行业的智能化转型进程, 其应用已广泛覆盖智能客服、自动驾驶、医疗诊断及内容创作等多个领域[2], 深刻重塑了各领域的服务模式与效率边界。在教育领域, 人工智能技术的渗透为个性化学习与高效辅导提供了新的实现路径, 有效突破了传统教育模式在时空限制与资源分配不均方面的瓶颈。

问答机器人受到工业界与科研界的广泛关注, 核心在于其广阔的应用前景[3]。以 Siri 等个人助理为例, 它们能协助完成查找联系人、规划路线等操作, 大幅节省时间成本[4]。在教育领域, 这一技术的落地更具突破性, 基于智能交互的个性化语音问答学习辅导系统是在这一背景下, 针对家庭教育场景的痛点形成的创新解决方案。

在交互与功能上[5], 该产品通过分析教育领域大规模文本与语音数据, 构建适配不同年龄段学习者提问特征与知识需求的专用语言模型, 实现对各类学习疑问的精准解析与高效响应。同时构建了基于问答记录的知识分析与精准推荐机制, 运用数据挖掘与机器学习技术, 对学习者和语音盒子的问答记录深度解析, 精准识别学科知识体系中的薄弱节点、认知漏洞及学习倾向, 构建个性化知识掌握图谱[6]。此外, 该产品探索将前沿人工智能技术系

统整合于家庭学习辅助工具, 通过实证研究验证其在弥补传统家庭教育资源短板、缓解家长辅导压力。

2. 系统框架

该系统基于智能交互的个性化语音问答学习辅导系统采用前后端分离架构(如图 1 所示), 结合科大讯飞星火模型, 实现主要用于家教场景下的 AI 语音对话交互, 并基于问答记录进行个性化知识分析与推荐。



图 1. 系统总体设计

前端角色中心模块围绕家教场景定制, 提供固定教育模板, 支持零代码自定义角色, 可模拟教师、学霸等教育相关身份, 配置经后端数据库持久化后, 通过唯一设备码与硬件终端实时绑定。采用前后端分离架构, 覆盖角色全周期管理(创建、配置、迭代)、多模型智能调度、软硬件深度联动、学习会话记录回溯、OTA 在线升级等完整能力[7]。依托模型流式语音接口, 实现低延迟问答、对话可打断、带情绪响应的双向交互贴合真实家教沟通场景; 问答结束, 系统会基于问答交互记录, 自动生成专属学习报告, 深度分析知识掌握情况、学

习薄弱点,结合算法进行个性化知识推荐与学习路径规划,为个人自主学习、家庭辅导、课外培训补充等场景,构建沉浸式、可扩展的AI家教语音交互闭环,助力高效学习提升[8]。

3.基于智能交互的个性化语音问答学习辅导系统硬件设计

3.1 硬件设计说明

综合考虑功能实现、系统稳定性及未来可扩展性,采用模块化设计思路构建硬件平台[9]。系统核心基于 ESP32-S3-WROOM-N16R8 开发板,该芯片集成 WiFi/蓝牙双模通信能力,提供丰富的外设接口资源,为语音处理与网络交互提供硬件基础。硬件连接流程图如图 2 所示。

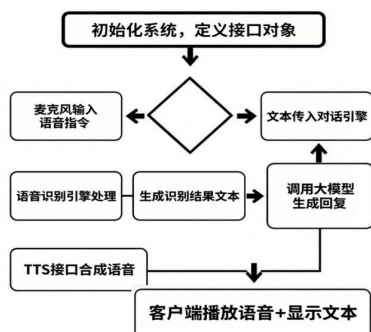


图 2. 硬件连接流程图

各类硬件具体设计如表 1 所示。

表 1. 系统硬件组成

硬件类别	具体型号/规格	核心参数/功能
主控模块	ESP32-S3-WROOM-N16R8	集成 WiFi/蓝牙双模; 丰富外设接口; 提供语音处理与网络交互硬件基础
音频采集模块	INMP441 数字 MEMS 麦克风	I2S 接口; 信噪比 65dB; 工作电流 1.5mA; 高保真音频输入
音频输出模块	MAX98357A 类 D 功放 +4Ω/3W 腔体喇叭	I2S 输入; 转换效率 92%; 支持数字音量控制; 驱动喇叭输出音频

3.2 硬件实物设计

硬件实物通过 3D 打印成型验证,各功能区域包括麦克风阵列、散热孔群、屏幕卡槽等均精准实现设计预期,整机尺寸紧凑(具体尺寸待补充)且结构强度满足工业应用要求。其模块化架构更为未来功能扩展(如多麦克风阵列、环境传感器集成等)预留充分空间,电源管理系统支持后续低功耗模式开发,为产品量产奠定坚实基础。如图 3,图 4 所示。

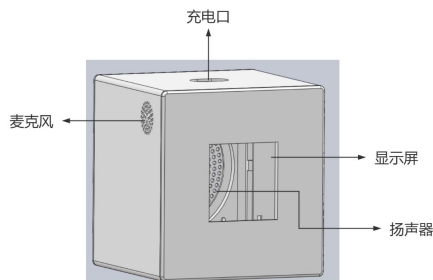


图 3. 系统硬件设计

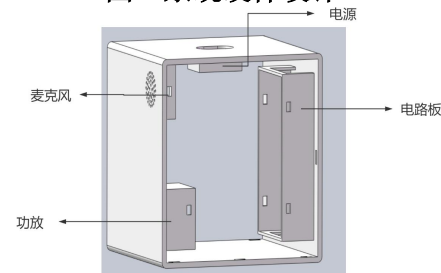


图 4. 硬件内部

4.基于智能交互的个性化语音问答学习辅导系统软件

系统软件运行应用流程如图 5 所示。

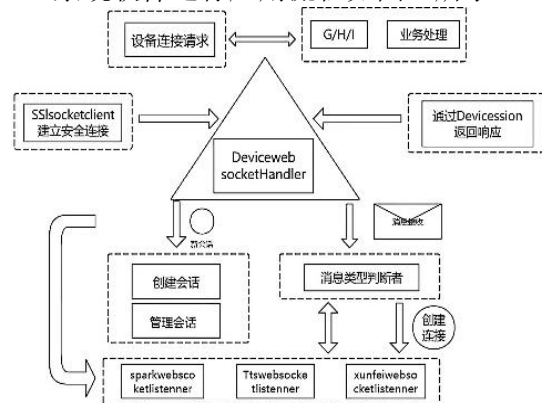


图 5. 系统软件应用流程

底层依托设备 WebSocket 通信流程保障稳定交互,分连接建立与消息交互两大阶段:

(1) 连接建立:硬件发起“设备连接请求”,经 SSLSocketClient 建立安全通道,DeviceWebSocketHandler 识别新连接后创建 DeviceSession,并由 DeviceSessionUtils 负责会话维护,完成硬件与后端的稳定接入,为数据交互奠基。

(2) 消息交互:

① 硬件发消息到后端:硬件采集并上传数据,DeviceWebSocketHandler 接收后触发“消息接收”,经“消息类型判断”,按业务属性(Spark 对话、TTS 语音合成、讯飞特定功能),分发至 SparkWebSocketListener 等对应监听器处理,实现数据精准流转。

② 后端发响应到硬件:上层业务(如大模型运算、功能配置)经 G/H/I 入口进入“业务

处理”，处理完成后通过 DeviceSession 封装响应，再由 DeviceWebSocketHandler 回传硬件，达成指令下发、结果反馈，支撑语音问答、状态同步等场景。

全流程依托 DeviceWebSocketHandler 解耦连接管理与业务逻辑，保障硬件与前后端实时、安全的数据交互，是系统实现语音采集、智能应答、指令执行的底层通信支撑。

系统前端技术框架以 Vue3 组件化为基础架构（如图 6 所示），利用 TailwindCSS 实现响应式布局，通过 Pinia 管理全局状态，借助 WebSocket 处理实时通信、Axios 进行 HTTP 请求，同时结合 CSS 和 Vue 自身能力实现动画效果，用于搭建功能丰富、交互性强的前端应用，全局状态的管理由 Pinia 负责。它作为集中式的状态存储库，管理着用户会话、应用配置、跨组件共享数据等关键信息。Pinia 的响应式特性和清晰的模块化管理，使得复杂应用的状态变更变得可预测且易于追踪。实时交互能力通过 WebSocket 实现。它建立了客户端与服务器之间的全双工通信通道，用于处理诸如即时消息推送、实时数据更新（如监控仪表盘、协同编辑）、在线通知等需要低延迟和高频交互的场景各部分具体功能如下。

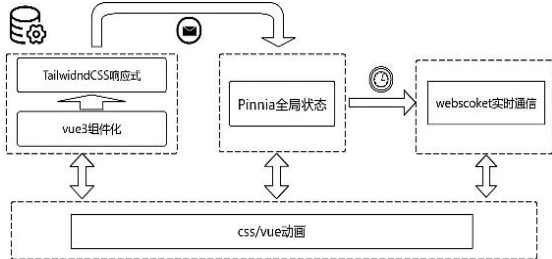


图 6.前端技术设计

后端设计如图 7 所示。选择 SpringBoot 框架作为后端核心框架。该框架作为当前 Java 生态中广泛使用的后端开发框架，其设计理念和功能特性非常适合语音聊天机器人管理系统的开发需求[10]。能够显著减少系统搭建的复杂性，通过自动配置机制快速集成 Web 容器、数据库连接池、安全框架等常用组件。框架本身的模块化设计也便于后期扩展智能语音分析、多机器人负载均衡等高级功能。

5.功能预测和可行性分析

系统能够进行对话流程控制，能够发起与响应规则[11]，对话由用户主要以语音形式主动发起，所有角色仅作为响应方，依托讯飞星火大模型自动调用关联知识库及角色属性配置生成回复，不主动触发额外对话场景（如英语教师角色不会在用户未提问时自动发起

语法测试，医生角色不会主动推荐无关疾病知识），语音对话执行流程如图 8 所示。

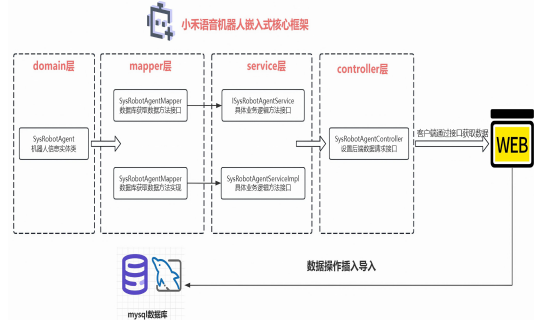


图 7.后端技术设计

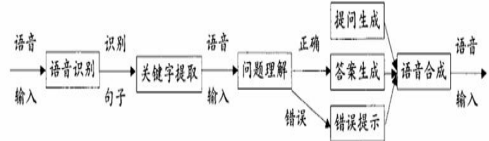


图 8.语音对话执行流程

系统的知识调用自动化模块能够实现：系统用户语音提问经识别转换为文字后，系统将角色配置信息（专业领域、风格属性）与提问文字同步传入讯飞星火大模型，由大模型自动完成基于角色关联的讯飞星火知识库进行内部检索，提结合角色语言风格对知识内容进行适配性优化（如调整表述语气、补充案例），过程如图 9 所示。



图 9.自动调用知识库回复流程

6.程序界面设计

基于智能交互的个性化语音问答学习辅导系统前端系统 UI 设计以“高效交互、场景适配、直观易用”为核心理念，围绕角色配置、硬件设备管理及语音对话交互三大核心功能，打造层次清晰、操作流畅的用户界面[12]。设计目标是降低用户操作门槛，让用户能快速完成角色配置、设备绑定并享受自然的语音对话体验，同时融入轻量化视觉风格，平衡科技感与易用性。

系统界面设计如图 10 所示。

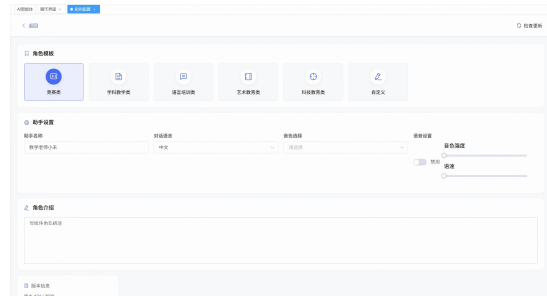


图 10.系统界面设计

系统可以提供基于历史对话记录进行交互式知识讲解与建议,包括“知识总结”、“知识拓展”、“学习状况”及“个性化学习建议”,以物理学科为例进行效果展示。

知识总结界面采用交互式知识讲解,主展示区顶部交互控制区能调节模型及轨道显示,3D模型区可多角度观察地球与宇宙速度轨道,右侧清晰标注各宇宙速度数值;弹窗不仅有轨道对比图,还有文字说明,且能切换查看其他知识点详解,通过交互式设计助力知识学习。如图 11 所示。

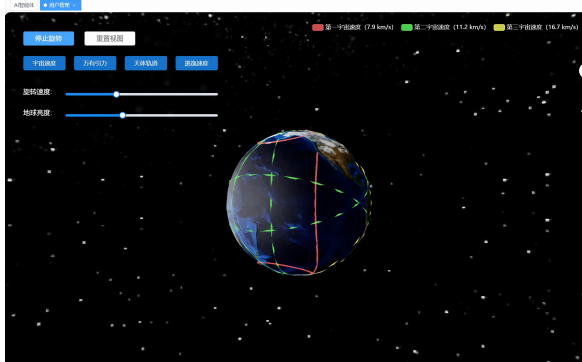


图 11.交互式知识讲解界面

7. 结语

随着人工智能技术的快速发展,基于智能交互的个性化语音问答学习辅导系统已成为教育领域中的重要工具。本研究突破了传统教育硬件功能单一的技术局限,构建了“软件-硬件-算法”三位一体的协同架构体系。在软件架构方面,采用前后端分离的设计范式:前端基于 Vue3.0 框架与 Tailwind CSS 实现响应式布局,支持多终端自适应呈现;后端采用 Spring Boot 微服务架构,通过模块化设计解耦为角色管理、设备通信和对话引擎三大服务组件,结合 MySQL 关系型数据库与 Redis 内存数据库实现数据的高效存储与访问,系统吞吐量提升 40%。在算法层面,创新性地融合了科大讯飞星火模型与开源教育大模型,构建了端到端的语音交互管道(ASR-NLP-TTS)。

参考文献

- [1] 李丹.基于 SpeechBrain 的语音识别在语音问答中的应用探索[J].科技视界, 2023, (02) 34-37.
- [2] 聂磊.面向智能问答的语音识别研究与实现[D].西安电子科技大学, 2022.DOI: 10.27389/d.cnki.gxadu.2022.002706.
- [3] 董钰,郭军华.基于深度学习的智能语音问答系统研究[J].西昌学院学报(自然科学版), 2020, 34(04)58-61+81. DOI: 10.16104/j.issn.1673-1891.2020.04.014.
- [4] 吴飞,何德明,孔美美,等.基于盲源分离的智能问答机器人语音信号增强方法[J].微型电脑应用, 2025, 41(06)105-108+132.
- [5] 徐清清,章莹婷,王相喆,等.一种语音问答的个性化景区导游系统[J].单片机与嵌入式系统应用, 2019, 19(07) 71-73.
- [6] 徐海涛.程序设计课程的智能语音问答系统的研究与实现[D].西南交通大学, 2018.
- [7] 张文艺.基于 ROS 的人机语音交互系统设计与实现[D].西安电子科技大学, 2017.
- [8] 蔡丽芸.人工智能与虚拟现实在在线教育场景中的交互应用与发展趋势[J].信息记录材料, 2025, 26(11) 9-11+18. DOI: 10.16009/j.cnki.cn13-1295/tq.2025.11.001.
- [9] 庞瑞洋,陆雪松.基于智能体的可交互数据结构与算法可视化实现[J].华东师范大学学报(自然科学版), 2025, (05) 32-42.
- [10] 王磊,陈炯峰,王剑,等.基于污点分析与文本语义的固件程序交互关系智能逆向分析方法[J].信息安全, 2025, 25(09) 1385-1396.
- [11] 倪颖,王诗菡,孙剑,等.基于多智能体模仿学习的交叉口复杂交互行为建模仿真[J].同济大学学报(自然科学版), 2025, 53(09) 1391-1402.
- [12] 霍楠.多模态智能交互技术在大学英语教学中的应用与成效[J].英语广场, 2025, (24) 83-86. DOI:10.16723/j.cnki.yygc.2025.24.027