

✦ 通慧智教 ✦

大模型辅助教学



优质平台 助你创造不凡

适用对象

学生、老师、平台管理者

主要功能

相似例题生成、薄弱例题推荐、学生作业评估、学习计划生成、辅助理解题目与答案评估

AI 赋能 · 【通慧智教】

项目详细方案

队名：CoooooooooI

组长：李昂

组员：许祖耀 吕栋 王清恒 蔺洋

目录

一、前言.....	4
二、创意描述.....	4
三、功能介绍.....	5
3.1 VR 端	5
3.2 网页端	5
四、特色综述.....	10
4.1 大语言模型数据.....	10
4.2 个性化	10
4.4 训练大模型	12
五、系统分析.....	13
5.1 系统用例图	13
5.2 逻辑视图	13
5.3 开发视图	15
5.4 进程视图	16
5.5 物理视图	17
六、详细设计.....	17
6.1 算法设计	17
6.2 系统设计	23
七、项目原型.....	28
7.1 低保真模型	28
7.2 高保真模型	32
八、开发工具与技术.....	36
九、应用对象.....	37
9.1 网页端	37
9.2 VR 端	37
十、应用环境.....	38
十一、结语.....	38

一、前言

近年来，人工智能教育技术（AIGC）以其强大的潜力和应用前景，在教育领域引起了广泛的关注和探讨。AIGC 技术的蓬勃发展不仅带来了全新的生产力革新，而且正在深刻地影响着教育生态的变革。然而，随着 AIGC 技术的迅速普及和应用，如何正确地利用这一技术主导教学过程，成为了一个亟待解决的问题。

在这个背景下，通慧智教站在教育科技的前沿，积极探索并应用 AIGC 技术来改进教学质量和效果。通慧智教的方法包括但不限于精细调整的语言模型（LLM）、构建高质量的数据集、打破传统的交互方式，并采用先进的算法来助力个性化学习。这些举措旨在解决当下教育资源同质化严重、教学过程缺乏互动、个性化学习内容匮乏以及学习效果反馈不足等问题。

通慧智教主要有以下特点：

1、通慧智教通过精细调整的语言模型（LLM），能够更准确地理解学生的需求和问题，从而提供更个性化、精准的教学内容和解答。这种个性化的学习体验不仅能够激发学生的学习兴趣，还能够更好地满足不同学生的学习需求，从而提高整体的教学效果。

2、通慧智教注重构建高质量的数据集，以确保教学内容的准确性和有效性。通过收集和整理大量的教学数据，通慧智教能够不断优化和完善教学模型，从而提供更加全面、深入的教学服务。

3、通慧智教还致力于打破传统的教学交互方式，采用更加灵活、多样化的教学方式和工具。通过结合虚拟现实、增强现实等前沿技术，通慧智教能够为学生提供更加生动、直观的学习体验，从而提高学生的学习积极性和参与度。

4、通慧智教通过采用先进的算法和技术，实现了教育的个性化。通过分析学生的学习数据和行为模式，通慧智教能够为每个学生量身定制学习计划和内容，从而最大限度地发挥每个学生的潜能，提高学习效果。

通慧智教在利用 AIGC 技术推动教育革新方面取得了优秀的成果。通过精细调整的语言模型、高质量的数据集、打破传统的交互方式以及个性化的学习内容，通慧智教为解决当下教育面临的种种挑战提供了创新的解决方案，为构建更加健康、高效的教育生态作出了积极的贡献。

二、创意描述

通慧智教借助于先进的多维数据分析技术，精心构建了学习者的详细画像，并利用经过细致调优的大型语言模型（LLM）生成定制化的学习报告。这一创新方法不仅帮助用户深入理解各类题目，还能够自动生成针对性的类似习题，极大地丰富了学习资源和方法。我们的数据采集网络涵盖了分布式爬虫、智能文档处理（IDP）技术以及各类开源数据集，通过精细的数据清洗流程，确保了信息的高质量 and 准确性。

在交互式学习体验方面，通慧智教推出了虚拟教学助理“小慧”及其 VR 课堂平台，通过结合语音到文本（STT）、大型语言模型（LLM）、文本到语音（TTS）技术以及其他

开发工具，为学习者提供了一个充满互动性和反馈的学习环境。此外，我们采用了一种精确的调优方法来优化本地 LLM 的性能，并使用 Rouge 得分系统来评估 LLM 的效率和准确性，确保我们的技术和服务能够满足用户的高标准需求。

三、功能介绍

3.1 VR 端

3.1.1 物体交互

用户在使用 VR 设备进入我们提供的虚拟现实环境时，将能够与多种趣味物体进行直观且互动性强的交互。这些物体设计得既有教育意义，又不乏娱乐性，这种交互方式增强了学习的趣味性，实现寓教于乐的教学理念。

3.1.2 网页端大屏

网页端全部功能迁移。通过 VR 设备，用户不仅能够进入一个全面沉浸的学习环境，还能在这一环境中通过网页端大屏进行学习。这种模式允许用户在虚拟现实中操作一个虚拟的大屏幕，就如同他们在现实生活中使用电脑或平板一样。用户可以浏览网页、观看教学视频等，所有这些活动都是通过虚拟现实中的简单手势和动作来完成。这不仅使学习变得更加直观和互动，还为用户提供了一个新颖的、没有物理限制的学习平台，极大地增强了学习的可访问性和便利性。

3.1.3 虚拟教学助理

为了提升用户在 VR 学习环境中的体验，我们特别设计了虚拟教学助理“小慧”。小慧是一个智能的虚拟角色，能够时刻陪伴在用户身旁，提供实时的学习辅导和帮助。无论用户遇到任何学习上的问题，小慧都能够提供及时帮助。

3.2 网页端

3.2.1 注册

用户可以使用学号来获取属于自己的账号，账号关联了这个用户的全部信息。

3.2.2 登录

用户可以通过设置的账号密码来登陆系统，进而使用系统所提供的其它功能。

3.2.3 编辑个人资料

用户登入系统之后可以在用户信息界面编辑个人的基本信息，如修改昵称、密码以及认证信息等。

3.2.4 注销账号

用户可以在用户中心界面将账号注销，数据库中也会清除与用户相关的信息。

3.2.5 虚拟教学助理

1. 语音转文本

通过整合百度的先进短语音识别 API，我们能够将用户的语音输入实时转换为文本，进一步输入给大型语言模型（LLM）。这一流畅的转换过程确保了用户可以无缝接收到 LLM 的智能回复，提升了交互的自然度和效率。

2. 文本转语音

在获得 LLM 的智能回复后，我们采用了本地预训练并微调过的 GPT-SoVITS 模型，将文本信息转化为语音信息。这段语音随后被后端传递至虚拟教学助理“小慧”，由她负责将这些语音反馈给用户，营造了一个互动且人性化的学习环境。

3. 口型同步

利用 uLipSync 技术，我们将虚拟人物的语音数据拆分为五个基本的元音音素，并为每种音素配备相应的模型口型。通过动态调整这五种音素的相似度，实现了虚拟人物口型与语音的实时精准同步，增强了虚拟教学助理的真实感。

4. 表情和动作控制

通过对大语言模型回复格式的精心设计，我们实现了直接通过 LLM 控制虚拟教学助理“小慧”的表情和动作。我们记录并将控制脸部表情的模型状态变化转换为动画片段，利用 Unity 的动画控制器进行精准控制；同时，借助 Unity 生成的 Humanoid Armature 和相应的动作动画，我们能够实现“小慧”在回复时进行复杂动作的展示。

5. 智能路由跳转

通过预设关键词和利用 prompt 工程技术改进 LLM 的回复格式，我们的系统不仅能够将回复发送至后端，还能够将其直接路由至 Unity 前端。一旦系统检测到回复中含有特定关键词，就会触发相应的操作，例如，当用户提供账号信息时，“小慧”能够自动跳转至登录界面并帮助用户填充必要信息，极大地简化了用户操作流程，提升了用户体验。

3.2.6 温馨短语生成

系统自动调用先进的商用大语言模型 API，为用户生成富有温情和激励的短语。这些精心挑选的短语不仅能够改善用户的学习环境，营造出一个更加友好和鼓舞人心的氛围，还能显著增强用户的学习动力。无论是遇到学习上的挑战还是需要额外的动力去探索新知，这些温馨短语都能为用户提供及时的心理支持和正能量。

3.2.7 网页走马灯大屏

网站主页上嵌入了使用 ElementUI 开发的走马灯组件。通过这个走马灯大屏，用户可以轻松地浏览项目的基本信息、网站功能介绍、最新动态等内容。这种视觉吸引力强的展示方式不仅能够加深用户对网站的认识和理解，还能有效地提升用户的浏览体验，增强用户对网站内容的兴趣。

3.2.8 精选课程推荐

网站采用了先进的大语言模型技术，根据用户登录时提供的学习信息等数据构建个性化的学习者画像，使大语言模型能够精准地挑选出适合每位用户的课程。通过大语言模型的智能分析和混合过滤推荐算法，我们从系统资源中为用户推荐量身定制的课程，旨在最大化满足用户的学习需求和兴趣。用户加入推荐的课程通过点击“加入课程”即可将课程加入到“我的课程”中。

3.2.9 智能题目推荐与解答

1. 题目推荐

在“我的学堂”界面用户可以在“前往学习”选项中找到系统推荐的题目，题目是系统根据用户的学习者画像利用本地使用优良的数据集精调后的大语言模型生成并经过 DACF 算法为用户精心推荐的，是适合用户具有针对性的提高薄弱知识点的题目。

2. 重新推荐题目

用户点击“让小慧换一组推荐题目”即可重新推荐题目，用户可以根据自身需求选择题目学习。

3. 题目类型

每张题目卡片都有题目类型的选项，我们将题目类型分为“简答”、“课程”、“判断”和“选择”，其中除了课程外的题目都为本地大语言模型生成并动态展示在网站中的，标有“课程”的题目即为每个课程所布置的题目，是由教师进行发布的。

4. 题目解答

每张题目卡片上都有“查看”按钮，点击后可以进入题目的详细页面，主要包括题目的要求以及介绍等信息。在本页面可以点击“让小慧给点建议”按钮可以查看大语言模型关于本道题目大语言模型给出的一些拓展补充说明，同时点击下方的“让小慧给点

提醒”按钮可以获得大语言模型关于本道题目的解析，但并不会直接给出答案，而且解析程度的深浅是由用户的学习情况来决定的，对于学习水平较高的用户小慧不会对题目进行过多的解释，可以在提高用户积极性的同时不降低学习效果，增加学生对题目的思考。

5.用户回答评估

当用户完成题目，提交答案后可以查看大语言模型对用户提交答案的解析和得分，解析中主要包括大语言模型对题目的回答以及对用户答案的点评。大语言模型在回答中可以详细给出用户的得分点以及失分点，以及用户如何在下次碰到类似题目时得到更高的分数，帮助用户科学学习。

6.相似例题生成

当用户点入题目卡片时在上方会有相似例题生成选项，在这里用户可以点击按钮“让小慧生成更多相似的题目”来生成与本道题目相似的例题。主要是根据将用户所点击卡片的例题输入本地大语言模型，而后让其生成相应的题目，并且对生成的题目大语言模型会评估其与例题的相似度，我们将相似度以图的形式展现在页面上，使用户有选择性地选择相似度高的例题进行学习。

3.2.10 学习进度及活动分析

1.课程学习进度

在课程学习进度中可以看到用户的课程学习进度条以及学习进度圆环，使用 e-charts 图标进行展示增强了数据的显示效果，便于用户自己把握相应的学习进度。

2.学习进度评估

我们将用户的课程学习进度暴露给商用大语言模型，使其完成对用户学习进度的评估，并将结果放在大语言模型回复框中，以文本的形式展现给用户，便于用户根据大语言模型的点评精准把控以及动态规划学习进度。对于不同用户的学习进度大模型会有不同的评价，助力实现定制化、个性化学习。

3.学习活动

我们记录用户的学习活动并将其以折线图的形式展现在网页上面，使用 e-charts 绘图，图表的横坐标轴为时间，纵坐标为用户学习课程的学习总时长。便于用户对自己的学习活动精准把握，规划学习时间。

4.学习活动评估

我们将用户的学习活动暴露给商用大语言模型，使其完成对用户学习活动的评估，并将结果放在大语言模型回复框中，以文本的形式展现给用户，便于用户根据大语言模型的点评精准把控以及动态规划学习时间。对于不同用户的学习进度大模型会有不同的评价，助力实现定制化、个性化学习。

3.2.11 个性化学习报告

1.学习报告生成

系统将用户在网站的学习信息等构建出学习者画像，并将其暴露给大语言模型，大语言模型对系统提交的数据（主要为用户的学习者画像）进行分析并生成独属于用户的学习报告，用户可以根据大语言模型生成的学习报告动态规划自己的学习，提高学习效率。

2.报告生成评分

当系统生成报告后用户可以对系统生成报告的真实性进行打分和评价，网站可以将评分暴露给大语言模型，以便可以生成质量更高的报告。

3.报告重新生成

用户点击按钮“让小慧重新生成一份”即可重新获得大语言模型的回复，重新生成定制化学习报告。

4.报告导出

用户点击“导出报告”按钮，即可导出屏幕中显示的报告，格式为 pdf。

3.2.12 做题历史保存

用户可以在做题历史选项卡中查看自己做过的所有题目，并且可以点击进入详细的题目界面。

3.2.13 多样化 UI 界面

在最上方选项卡中，用户可以根据自身需求自动切换前端主题界面，主要有白天和夜晚两种主题，使用户可以选择最适合自己的界面，符合用户喜好。

3.2.14 课程学习

1.课程简介

在课程学习中，点击即可查看课程简介，主要为课程相关学习内容、章节等信息。

2.课程视频

用户可以在课程学习中上方选项卡点击课程视频进入课程视频的学习界面，同时可以点击不同的章节进行学习视频的切换。

3.课程 PPT

教师可以将课程 PPT 上传至网站，同时用户可以点击课程 PPT 选项卡查看教师提供

的课程 PPT。

4.课程文档

教师可以将课程文档上传至网站，同时用户可以点击课程文档选项卡查看教师提供的课程文档。

5.问答讨论

在课程学习中，点击问答讨论可以进入问答讨论区，用户可以选择自己感兴趣的问答进入回答问题，同时对于每一个问题小慧给回给出 AI 回答答案，助力用户学习。

四、特色综述

4.1 大语言模型数据

本系统为大语言模型的训练提供了兼具质量与数量的文本数据。

4.1.1 数据来源

数据来源多元，涵盖了 scrapy-redis 分布式爬虫、智能文档处理以及协作平台等多个渠道。这些渠道提供了广泛而丰富的信息资源，包括网络上的数据、各种文档和协作平台的内容。通过这些多源数据，能够获取不同领域、不同形式的题库文本信息，从而为分析和应用提供了更全面和多样化的基础。

4.1.2 数据预处理

在数据预处理阶段，采取了一系列严谨的步骤以确保数据质量。首先，清除了文本中的无意义字符和停用词，这样可以消除对后续分析的干扰。其次，进行了噪声过滤，将不相关的信息从数据中剔除，确保处理得到的是高质量的数据。接下来，运用 LDA 主题分析技术对数据进行分类。最后，构建了多种形式的 FAQ 问答对，为用户提供更多样化的信息呈现方式，从而提高数据的可用性和用户体验。通过这些预处理步骤，系统能够更有效地利用数据，并为后续的分析和应用奠定了坚实的基础。

4.2 个性化

当前教育的一大痛点是教学同质化，学生得不到个性化教学。该系统从多个角度创新型地对学生进行个性化教育。

4.2.1 混合过滤算法推荐例题

通过混合过滤算法，系统综合考虑学生的学习历史、兴趣和知识背景，为每位学生提供个性化的推荐。这样的推荐系统旨在提高学习效率和成效，帮助学生针对性地加强

薄弱环节，更好地掌握知识。通过不断优化推荐算法，目标是为学生提供更加精准和有效的学习辅助，促进他们在学习过程中取得更好的成绩。

4.2.2 多维用户画像构建

通过分析学生的学习行为、成绩历史和兴趣偏好等多维数据，系统构建详细的学习者画像。这一画像包含了学生的学习习惯、强项和薄弱领域等信息，为个性化学习提供了基础。进一步利用大型模型技术，系统生成针对性的个性化学习报告。这些报告根据学生的学习画像，提供定制的学习建议、课程推荐和习题练习，帮助学生更高效地学习和提高成绩。这种个性化的学习方式能够更好地满足学生的需求，促进其学习动力和成长。

4.2.3 大模型辅助理解题目与答案评估

大语言模型在辅助理解题目与答案评估方面发挥着至关重要的作用。它不仅能够深入分析题目的内容和结构，还能够理解答案的逻辑和正确性。通过对大量数据的训练，这些模型已经能够辨识各种题型，从简单的选择题到复杂的解答题，甚至是开放式问题。此外，它们还可以评估答案的质量，包括逻辑性、相关性和完整性。更进一步的，大语言模型能够提供针对答案的反馈，指出答案中的错误或不足，并给出改进的建议。这一过程不仅加深了学生对问题的理解，也促进了批判性思维和创造性解决问题的能力。因此，大语言模型成为了教育和评估领域的一个宝贵工具，帮助人们更有效地学习和评估知识。

4.3 交互方式

针对交互方式与反馈手段的缺乏，系统采取了多种交互方向供学生畅游知识海洋。

4.3.1 虚拟教学助理

在构建虚拟教学助理"小慧"的过程中，采用了 Unity、Blender、VRoid 等先进工具和技术，以打造一个具有高度真实感的 3D 虚拟形象。这一虚拟助理不仅拥有生动的外观，还能够通过自然语言进行交互，极大地增强了用户体验。借助语音转文本 (STT) 技术，"小慧"能够准确地将用户的语音输入转化为文本信息，进而通过大语言模型 (LLM) 分析这些信息，理解用户的查询意图并提供相应的信息或解答。在回应用户时，"小慧"还利用了文本转语音 (TTS) 技术，将文本信息转换为流畅自然的语音输出。这种结合了视觉、听觉和语言处理的多模态交互方式，使得"小慧"不仅仅是一个虚拟助理，更是一个能够提供富有吸引力的学习体验和互动环境的平台。通过这种创新性的自然语言交互式平台，"小慧"能够在教育、培训等多个领域发挥重要作用，为用户提供个性化、高效率的学习和辅导体验。

4.3.2 VR 虚拟现实

VR 虚拟现实和 MR 混合现实在教学领域的应用正变得越来越广泛，为学习提供了前所未有的沉浸式体验。通过 VR 技术，学生可以进入完全由计算机生成的三维虚拟环境，不受物理空间限制，从而实现对复杂概念的直观理解和实践操作。与此同时，MR 技术通

过将虚拟物体叠加到现实世界中，使得学生能够以全新的方式与教学内容互动，弥合了虚拟与现实之间的界限。这些技术不仅增强了教学的趣味性和互动性，还大大提高了学习效率和记忆深度。通过这些先进技术，教育者能够创造出更加丰富多彩、高度个性化的教学环境。

4.3.3 本地 GPT-SoVITS 大模型

本地部署的 GPT-SoVITS 大模型，为语音合成领域带来了突破性的进展。这一技术使得学生能够以前所未有的速度和灵活性，从多种多样的声音中进行选择，以进行交互。不论是模拟特定的情绪，还是切换不同的语言和口音，GPT-SoVITS 都能提供高度逼真的语音输出。这种能力极大地丰富了教学互动，让学习变得更加生动有趣。学生可以通过与各种虚拟角色的互动，提高语言学习的兴趣，加深对知识点的理解。

4.4 训练大模型

系统使用了本地部署的大模型与云端的商用大模型。

4.4.1 本地大模型

采用多种大模型并通过不同的精调方式，如 FT (Fine-Tuning)、GaLore、LoRA (Low-Rank Adaptation)、以及 QLoRA，成功实现了相似例题生成的功能。这种方法允许模型在保持原有知识结构的同时，灵活适应特定的教学和练习需求。通过这样的技术组合，模型能够根据已有的题目，创造出具有相似难度和主题的新题目，为学生提供更加丰富和多样化的学习材料。这不仅增强了学习的深度和广度，还提高了学生解决问题的能力，特别是在准备考试或加强特定知识点理解方面显示出了显著效果。此外，这种方法还支持个性化学习路径的创建，能够根据学生的学习进度和理解水平，提供最适合的练习题，进一步提升学习效率。

4.4.2 商用大模型

经过训练和精调的商业大模型具备了在教育领域多方面应用的能力。它能够自动评估学生的作业，提供详尽的反馈和改进建议，帮助学生识别并强化弱点。同时，该模型还能根据学生的学习进度和理解程度，制定个性化的学习计划，确保学习效率最大化。此外，它支持自然语言的对话交互，能够回答学生的疑问，引导学习方向，增加学习的互动性。该模型还能控制虚拟人物，为学生提供模拟情境学习的机会，以及管理网站路由，引导学生高效地访问学习资源。这些功能共同为教育领域带来了革命性的变革，极大地提高了教育的可接入性、个性化和互动性。

五、系统分析

5.1 系统用例图

网页端用例图

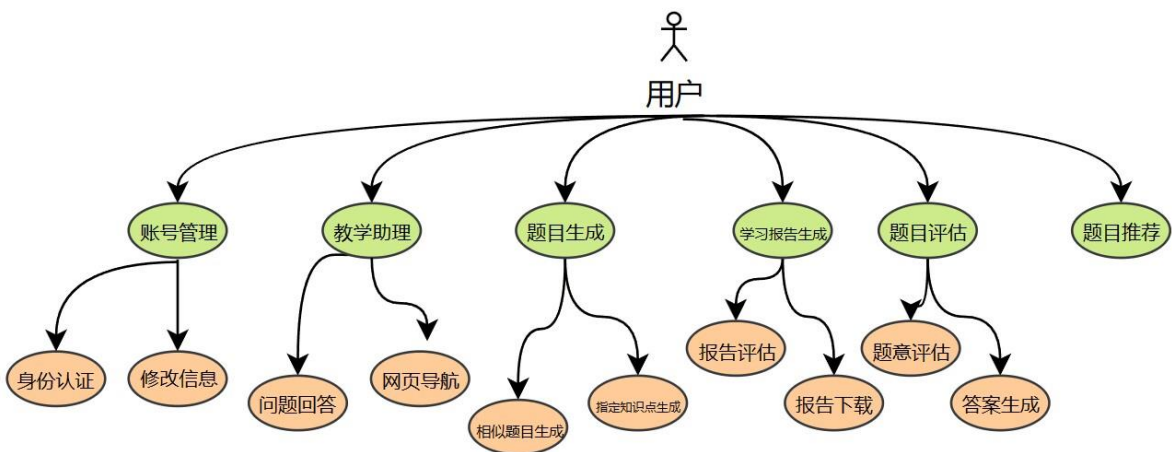


图 5.1.1 网页用例图

5.2 逻辑视图

逻辑视图描述通慧智教系统的功能需求，即通慧智教系统给用户提供服务，描述系统软件功能拆解后的组件关系、组建约束和边界，反映系统整体组成与系统如何构建的过程。

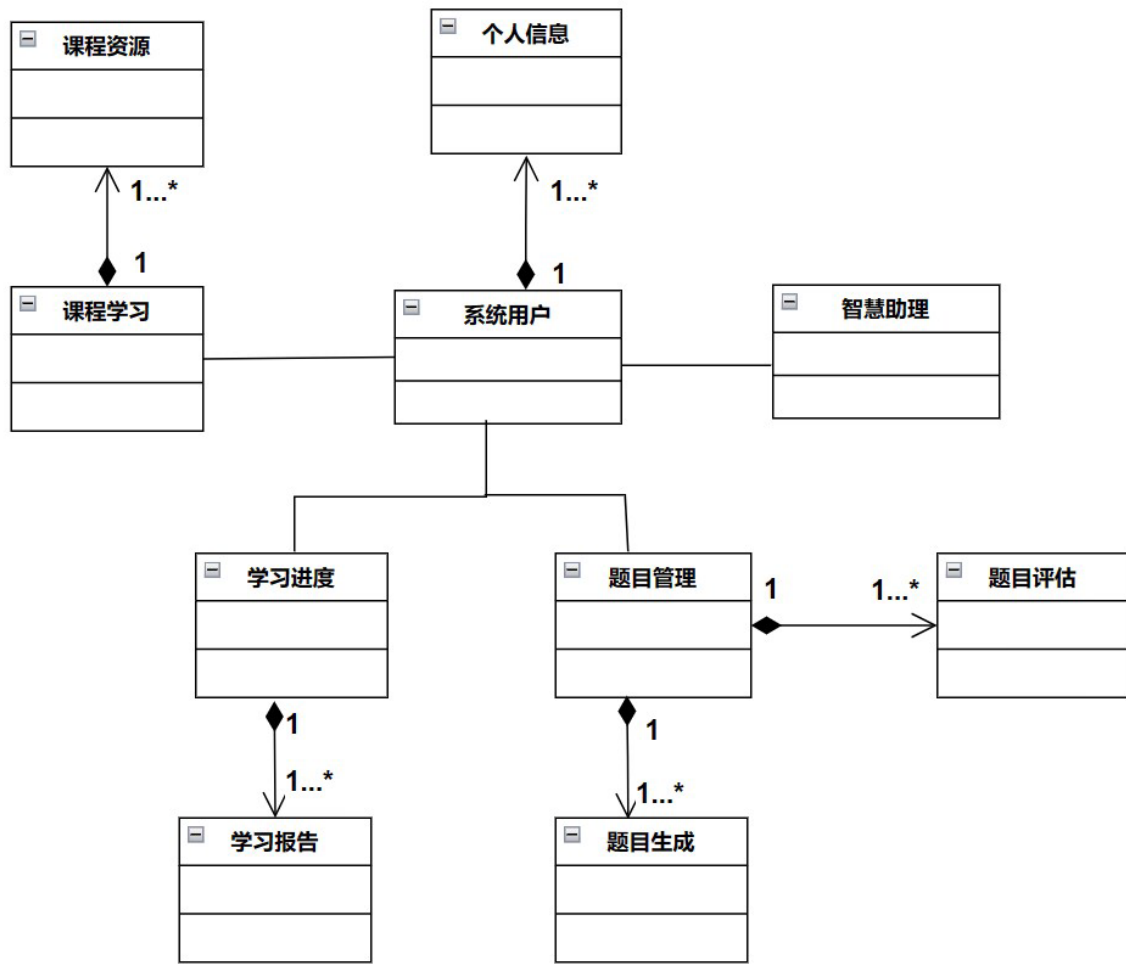


图 5.2.1 逻辑视图

5.3 开发视图

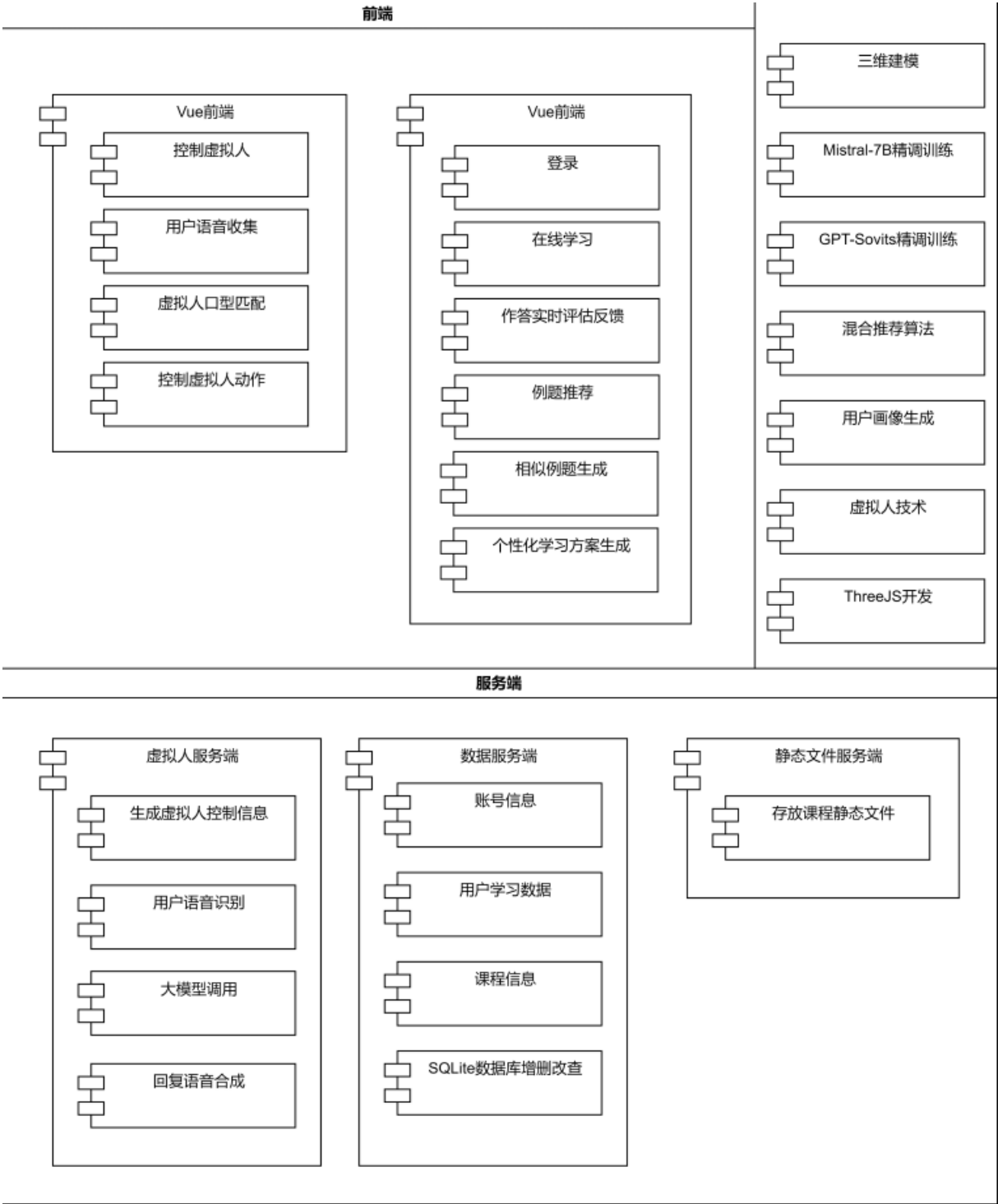


图 5.3.1 开发视图

5.4 进程视图

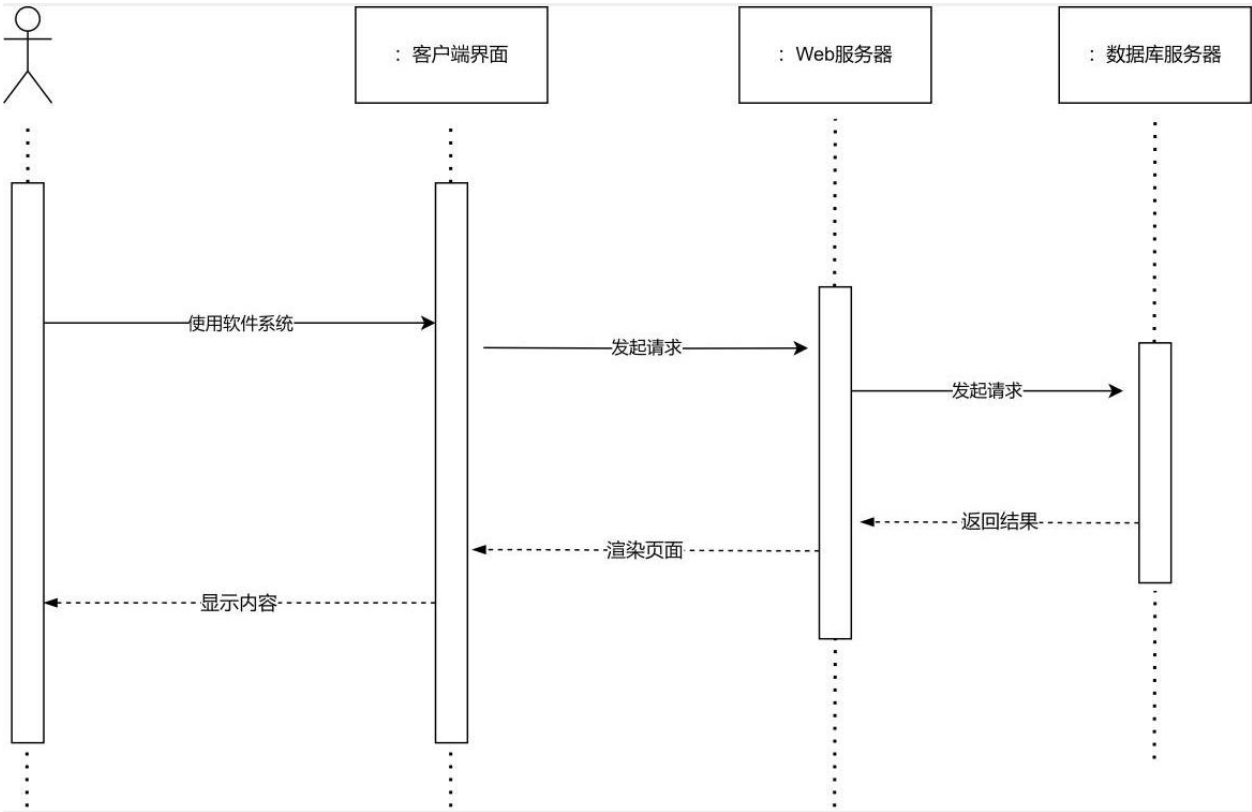


图 5.4.1 逻辑视图

5.5 物理视图

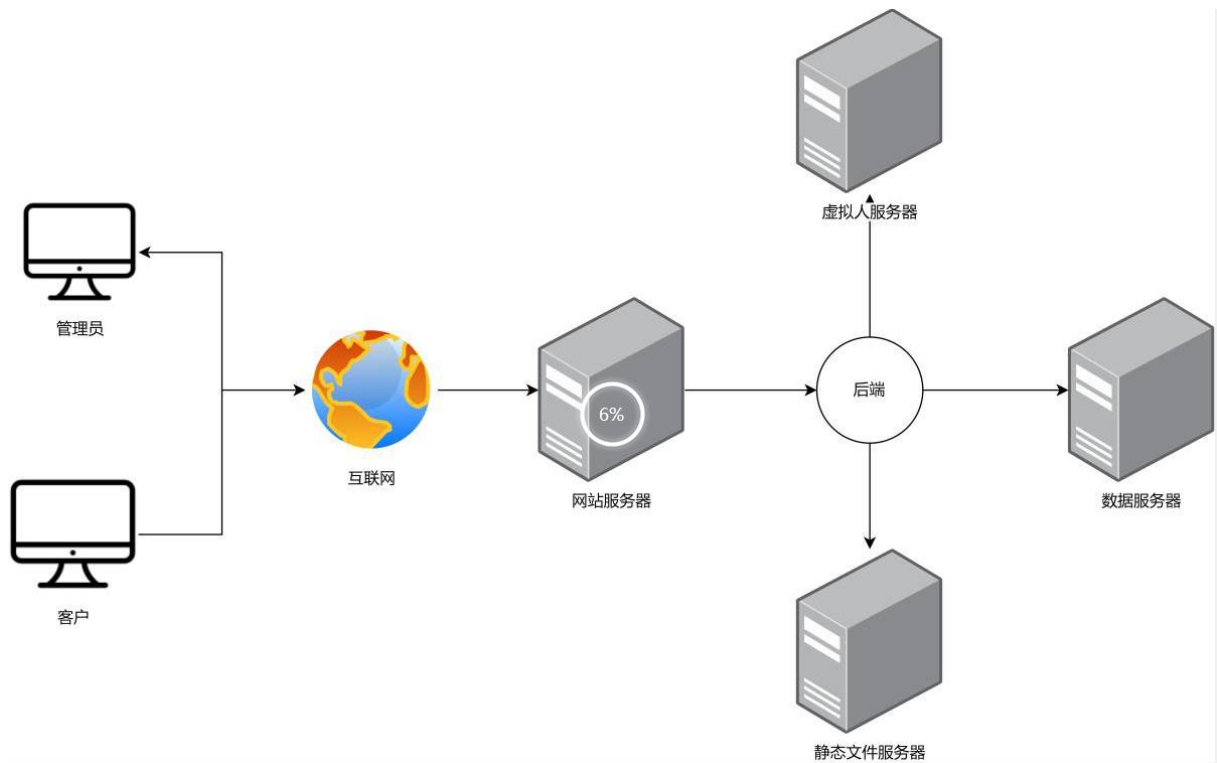


图 5.5.1 物理视图

六、详细设计

6.1 算法设计

6.1.1 学习者画像构建

1.学习者数据获取和处理

我们关注学习者在平台上产生的学习进度、学习偏好及历史成绩三个维度的信息，主要范围如下：

学习进度：包括用户完成课程的速度、课程完成度、特定时间内的学习频率等。

学习偏好：涵盖用户对不同学习内容（如视频、文章、练习题）的偏好、学习时长和频次、课程主题偏好等。

历史成绩：包括用户在平台上的测试成绩、课后练习得分、课程评价等

当获取学习者三个维度的信息后，我们清除了数据中的无意义字符和停用词，这样可以消除对后续分析的干扰。其次，进行了噪声过滤，将不相关的信息从数据中剔除，确保处理得到的是高质量的学习者数据。

2. 学习者画像构建

我们通过分析经过数据清洗后的学习者数据，提取其中有价值的信息，形成学习者的基本信息、内容偏好、学习风格等多维度特征标签：

基本信息维度：涵盖学习者的个人信息、学习背景等静态数据。

内容偏好维度：反映学习者对不同学习内容的偏好，包括课程类型、学习资源类型等。

学习风格维度：根据学习者在学习过程中的行为特征，如互动、讨论参与度、学习资源使用习惯等，划分学习者的学习风格

通过这三个维度，构建用户的学习者画像。

6.1.2 混合过滤推荐算法设计

我们采用基于注意力机制结合协同过滤模型构建高效的混合过滤推荐算法。

注意力机制：注意力机制（Attention Mechanism）是一种在深度学习模型中模拟人类注意力聚焦特性的技术，它使模型能够在处理大量输入信息时自动地将焦点放在最重要的部分。

注意力机制的计算公式如下：

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

其中 Q 代表 Query，表示当前我们想要提取信息的部分； K 代表 Key，表示输入数据中每个部分的标识符或索引，用于与 Query 进行匹配； V 代表 Value，表示输入数据中的实际内容，当确定了哪些 Key 与 Query 最相关后，相应的 Value 将被用来计算最终的输出或加权表示。

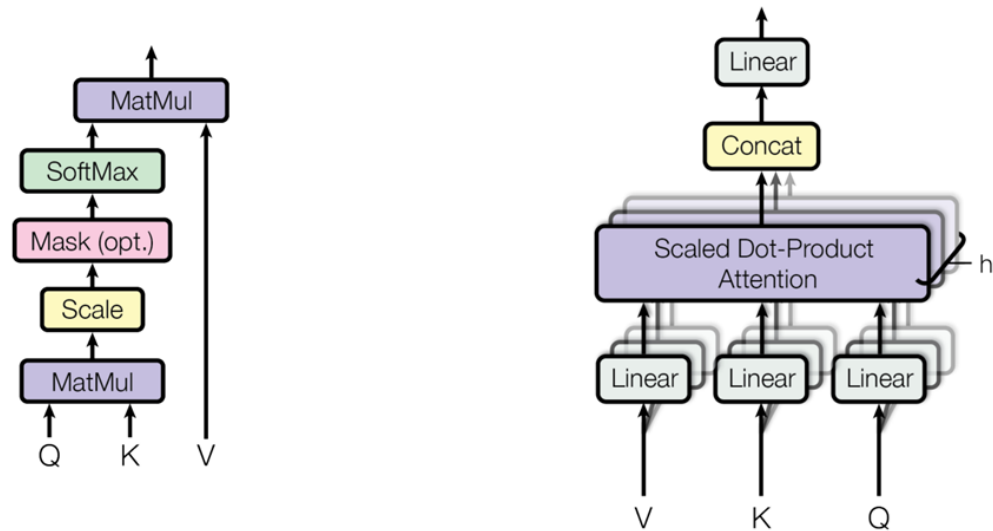


图 6.1.2.1 注意力机制与多头注意力

协同过滤：协同过滤算法是一种常用于构建推荐系统的技术，它通过分析用户对物品的喜好信息，预测用户可能感兴趣的其他物品。基本思想是，如果用户 A 和用户 B 在过去对某些物品的喜好上有相似之处，那么用户 A 可能会对用户 B 喜欢的其他物品感兴趣，反之亦然。主要分为基于用户的协同过滤和基于物品的协同过滤。

我们的混合过滤算法主要包括项目级的注意力机制和特征级的注意力机制的构建。

1. 项目级的注意力机制

项目级的注意力机制旨在学习用户对不同项目（如电影、商品等）的关注程度，并据此调整项目的重要性权重，以改善推荐的准确性和个性化。主要有以下三个流程：

（1）输入与权重学习：首先，算法会根据用户历史上的交互来获取隐性反馈。然后，通过考虑目标项目与用户历史交互项目之间的关系，使用多层感知器（MLP）学习每个历史项目的注意力权重。这一步的目的是为了允许不同的历史项目对用户当前偏好的表示做出不同贡献。

（2）权重分配：通过 Softmax 函数，将学习到的注意力权重转换为概率分布形式，这样可以确保所有历史项目的权重加起来等于 1，每个项目的权重反映了其相对于其他项目的重要性。

（3）用户表示与预测：最终，利用这些权重来加权汇总历史项目的特征，形成用户的表示向量。基于此向量，模型预测用户对未见项目的偏好，从而生成推荐列表。

2. 特征级的注意力机制

特征级的注意力机制着眼于学习每个项目内部的特征（如题目的长短等）对用户偏好的贡献度，以便更细致地捕捉用户的个性化兴趣。主要有以下三个流程：

（1）特征提取：对于每个项目，提取其特征并将它们表示为向量形式。这些特征向量可能包含了该项目的多个属性信息，例如题目的类型、长短等。

（2）学习特征权重：利用 MLP 模型，以目标项目的特征和用户历史偏好项目的特征作为输入，学习每个特征对于用户偏好的权重。这一过程允许模型区分同一项目内不同特征对用户影响的大小，反映用户对项目不同方面的关注度。

（3）加权特征向量生成：基于学习到的特征级注意力权重，将各个特征向量加权汇总，生成反映用户偏好的加权特征向量。

我们综合考虑两级注意力，将项目级和特征级注意力机制结合起来，通过双重调整，精确地捕捉用户的个性化偏好，从而实现对用户感兴趣的题目更准确、更个性化的推荐。

6.1.3 智能文档处理与分类

1. 文档收集与数据提取

为提高含图像 PDF 文档（如王道计算机考研 408 等）的文本识别准确性，首先进行图像预处理，包括去噪、旋转校正和亮度调整。这些步骤有助于清晰化图像中的文本，

减少识别错误。

接着，利用 OCR 和 LaTeX 识别算法，我们能够高效地从以图像为载体的习题集中提取数据。OCR 负责文本识别，而 LaTeX 识别专注于数学公式和专业符号，确保从图像中提取的数据不仅准确，还能保持原有的格式和意义。在提取过程中，特别注意分析题目的完整性和语义的正确性，避免因图像质量问题如模糊或遗漏导致的信息丢失。随后，通过设定的业务规则和数据模型，对提取出的信息进行严格验证，确保每一条数据都是完整且准确的。这一流程不仅提高了数据处理的效率，还大大降低了人为错误，保证了数据的高质量和可靠性。

2. LDA 分类

LDA 可以看作是两个过程，即 LDA 作为生成过程、LDA 作为推理过程。当我们想要创建语料库（一组文档）时，LDA 作为生成过程实现，其中我们已经知道每个主题的单词分布值(φ_k)以及每个文档的主题(θ_d)比例。相反，当我们想要识别潜在变量时，LDA 作为推理过程，包括每个主题的单词分布和每个文档的主题比例，并且我们只有一组单词作为观察变量。

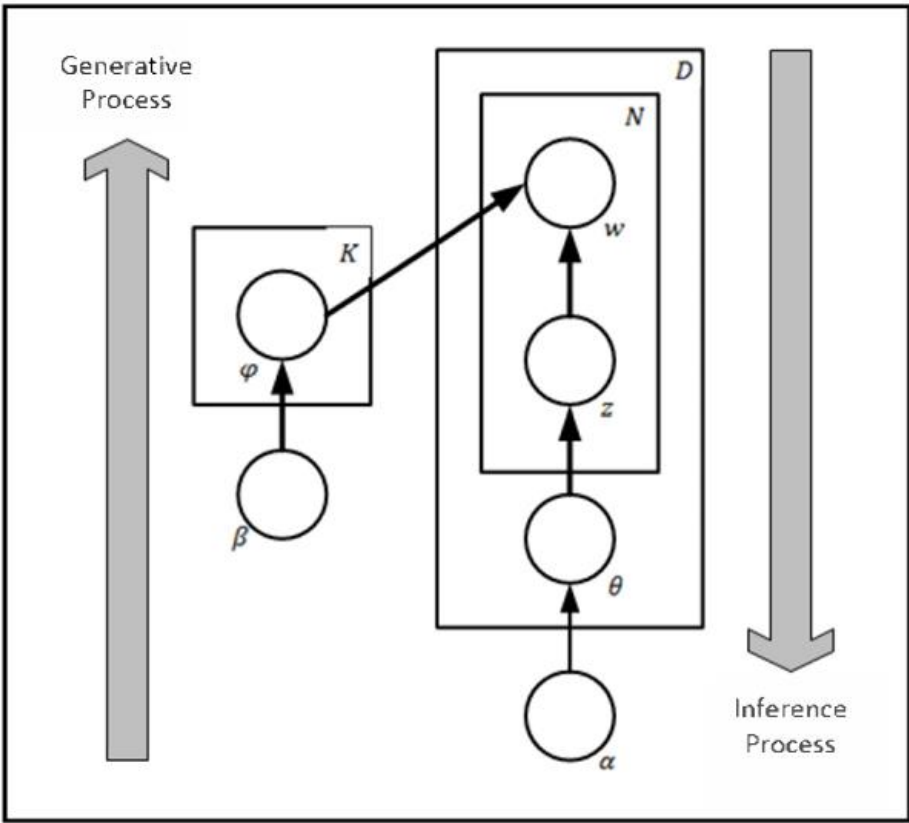


图 6.1.3.1 潜在狄利克雷分配的图形模型

在利用线性判别分析 (LDA) 进行分类任务时，整个流程分为训练和测试两个阶段，涉及到两类数据：训练数据和测试数据。为确保模型的泛化能力，采用 5 倍交叉验证策略对数据集进行划分和评估。无论是在训练阶段还是测试阶段，都会对文本进行预处理，这包括标记化（将文本分解为单词或短语）、删除停用词（去除常见但对主题不具有辨

识度的词汇)，以及词干提取（归纳单词到其根形式）。这些步骤有助于提炼出对分类任务更有意义的特征，进而提升 LDA 模型的性能。

在测试阶段，我们通过 Kullback-Leibler Divergence (KLD) 来评估测试文档的主题分布与训练阶段确定的类别主题分布之间的相似度。KLD 值越小，表示分布越相似。这种方法的测试结果显示，模型的最佳整体精度大约达到了 70%，这意味着我们的分类模型能够以较高的准确率将文档归类到其对应的主题中。

6.1.4 大语言模型精调与评估

在此模块，我们采用精调评估后的本地部署大模型 **Mistral-7B** 实现相似题目生成的功能，采用商用大模型-文心大模型 **ERNIE 3.5** 实现作业评估、生成个性化报告的功能以及采用精调后的 **ERNIE Speed** 为虚拟教学助理生成互动文本。

1. 本地部署大模型精调评估

我们采用了 **LLaMA-Factory** 开源 LLMs 精调评估框架来对大语言模型进行精调和评估，训练并选取出了实现相似题目生成功能效果最好的大语言模型。

LLaMA Factory 是一个 LLM 微调工具，支持预训练，监督微调和奖励建模训练模式。每种模式都支持 LoRA 和 QLoRA 微调策略。常用的本地大语言模型有 ChatGLM、Mistral、Gemma、LLaMA 等。我们根据下载的模型选择了 LLaMA2-7B、Mistral-7B、Gemma-7B、Qwen1.5-7B、Yi-6B、ChatGLM3-6B 这 6 个开源大语言模型，我们使用这 6 个大语言模型分别通过 Baseline、FT、GaLore、LoRA、QLoRA 这 5 个精调方法进行精调与对比评估。

对于这 5 个精调方法，**Baseline** 是在进行实验或评估时作为比较标准的基本方法；**FT** 精调方法通常指在已经训练好的模型上微调参数，以提高性能或适应特定任务；**GaLore** 是一种基于遗传算法的超参数优化方法，通过模拟生物进化的过程来搜索最佳超参数组合；**LoRA** 微调技术使用低秩近似方法来减少将十亿参数模型适应特定任务或领域的计算和财务成本；**QLoRA** 是一种高效的大型语言模型微调方法，可以在保持完整的 16 位微调性能的同时显著减少内存使用。

我们通过多模型多精调方式对比在基准数据集上的 Rouge 得分来选取实现实现相似题目生成功能效果最好的大语言模型。其中 **Rouge** 是一种用于评估文本摘要（或其他自然语言处理任务）质量的指标，主要关注生成的摘要中是否捕捉到了参考摘要的信息，着重于涵盖参考摘要的内容和信息完整性。通过 Rouge 得分，我们可以评估大语言模型生成文本的信息完整性和涵盖程度。经过多次实验，**Mistral-7B** 通过 LoRA 精调获得了最高得分。所以我们选取了精调后的 **Mistral-7B** 完成模型相似题目生成的功能。

Model	Baseline	FT	GaLore	LoRA	QLoRA
LLaMA2-7B	12.94	22.87	22.40	22.70	22.61
Mistral-7B	14.39	22.03	22.99	23.47	23.28
Gemma-7B	15.97	22.07	/	22.41	22.44
Qwen1.5-7B	15.40	22.46	21.76	22.71	22.52
Yi-6B	16.85	22.40	22.68	22.98	22.97
ChatGLM3-6B	18.51	22.00	22.16	21.68	21.70

图 6.1.4.1 模型对比

2.商用大模型精调训练

对文心大模型 **ERNIE Speed** 进行精调训练，我们首先搜集了各平台上的课程相关信息，使用百度提供的 **FAQ 挖掘服务**，自动挖掘 **FAQ 问答对**，构建数据集、对模型进行训练。



图 6.1.4.2 商用大模型

我们通过千帆大模型平台提供的 **SFT 任务**，以 **ERNIE Speed** 为基准模型进行进一步的 **Fine-Tuning**，并将模型发布以供平台调用。

< 返回

创建SFT任务

基本信息

* 任务名称:

计算机网络原理

支持中英文、数字、下划线(_), 2-20个字符, 不能以下划线为开头

任务类型:

文本生成

任务描述:

请输入

0/500

选择基础模型:

ERNIE Speed

ERNIE Lite

ERNIE Tiny

Llama-2

ERNIE Speed为百度自主研发的文心产业级知识增强大语言模型（高性能版），相较ERNIE Lite实现了基础

模型版本:

ERNIE-Speed-8K

单条数据支持8192 tokens。模型基于海量高质量数据训练，具有更强的文本理解、内容创作、对话问答等能

训练配置

训练任务的算法选择、参数及相关配置，训练配置参数影响训练速度及模型效果。

增量训练:

☐

训练方法:

☐ 全量更新 ☒ LoRA

LoRA在固定预训练大模型本身的参数的基础上，在保留自注意力模块中原始权重矩阵的基础上，对权重矩阵支持发布私有资源池预测服务[查看资源池说明](#)

参数配置:

超参数	数值	说明
-----	----	----

确定

取消

预估费用:
¥7.62 ~ ¥12.19 限时5折 计费说明
¥15.24 ~ ¥24.38

图 6.1.4.3 精调大模型

6.2 系统设计

6.2.1 总体架构图

系统总体架构图如图 6.2.1.1 所示：

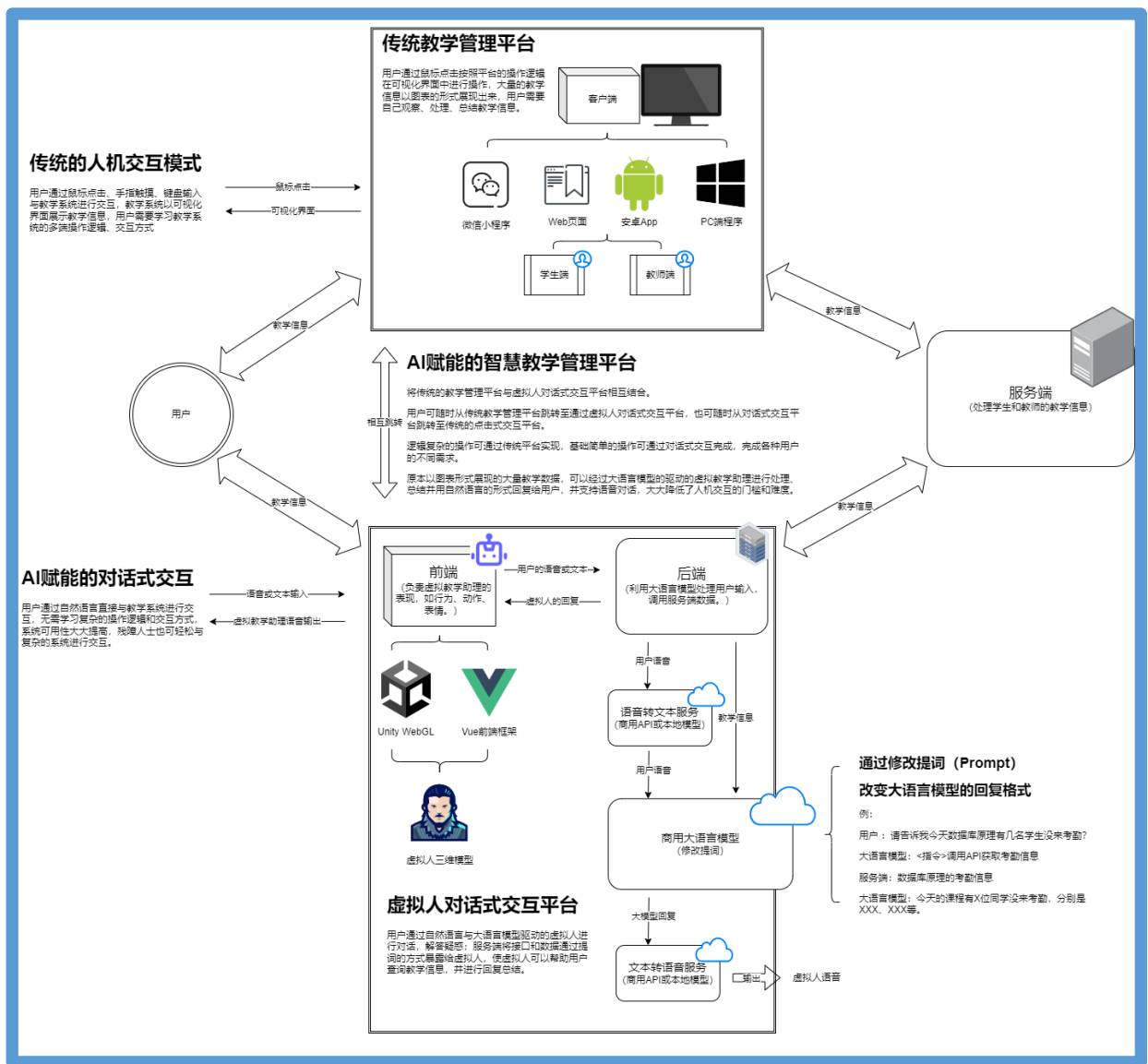


图 6.2.1.1 总体架构图

6.2.2 架构设计

平台开发采用了前后端分离的开发范式，前端负责页面呈现，虚拟人控制，后端则负责具体的业务操作、大模型相关技术、存储相关的静态文件等。

系统架构图如图 6.2.2.1 所示：

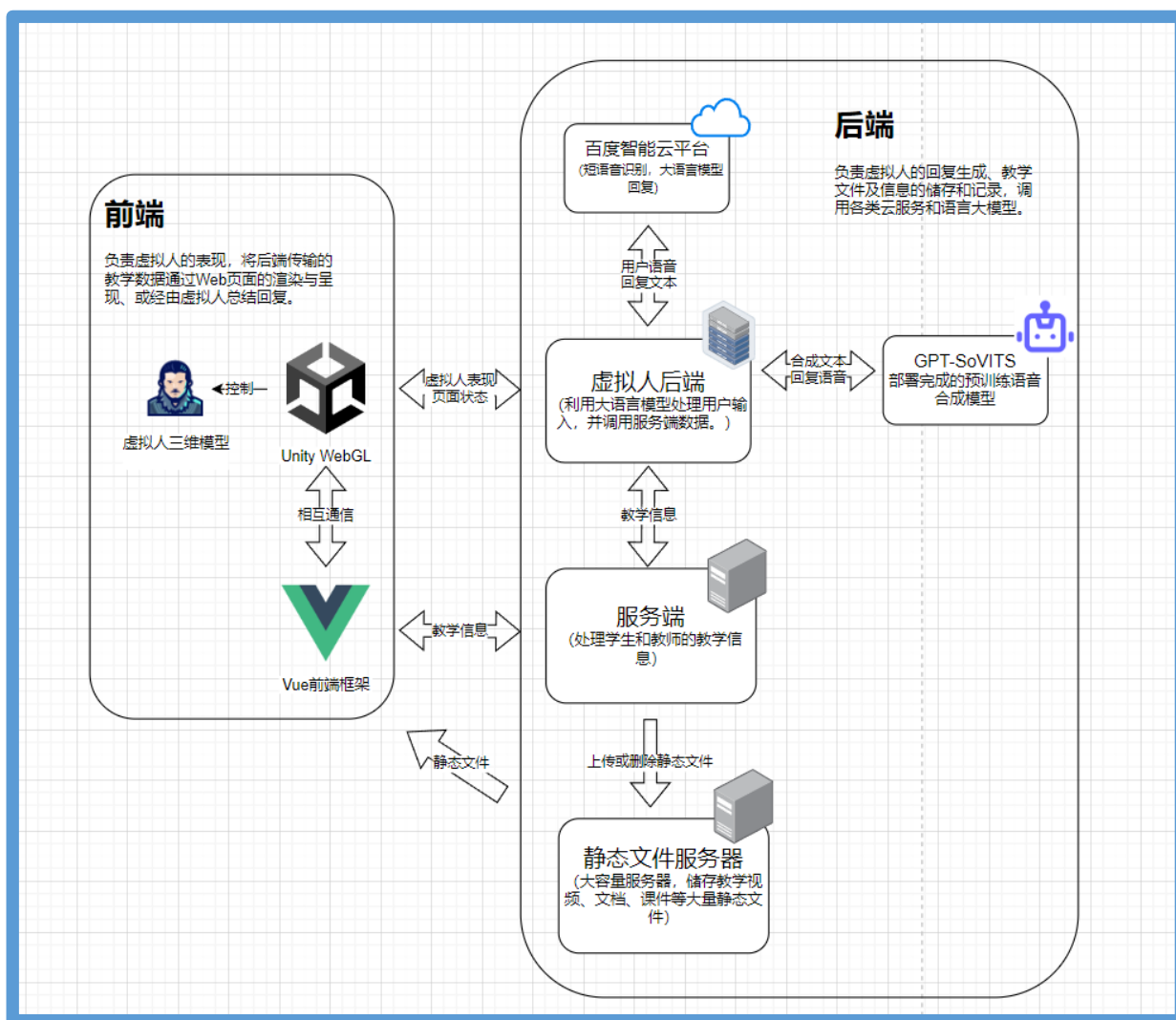


图 6.2.2.1 系统架构图

项目总计设计到两个前端、三个后端，一个云平台，两个本地预训练精调模型。

前端分为 Unity WebGL 程序和 Vue.js 前端程序，其中 Unity WebGL 程序负责控制虚拟人，Vue.js 前端程序负责页面呈现，两个前端程序相互通信，从而实现虚拟人对网页路由、功能的直接控制。

后端则分为三种，一是虚拟人后端，负责将用户的语音或文本输入进行语音识别或发送给大语言模型获得回复，再将大语言模型的回复进行语音合成，生成虚拟人语音并传输给 Unity WebGL 前端程序，使虚拟人进行表情、动作、口型的改变。

二是数据服务端，负责储存网站的业务数据，包括但不限于学生和教师的教学信息，课程的相关信息和内容等，三是静态文件服务器，静态文件服务器负责储存在网站教学过程中所需要的视频、PPT、文档、课件的静态文件，并在需要的时候传输给用户以进行下载。

6.2.3 系统架构

系统架构图如图 6.2.3 所示：

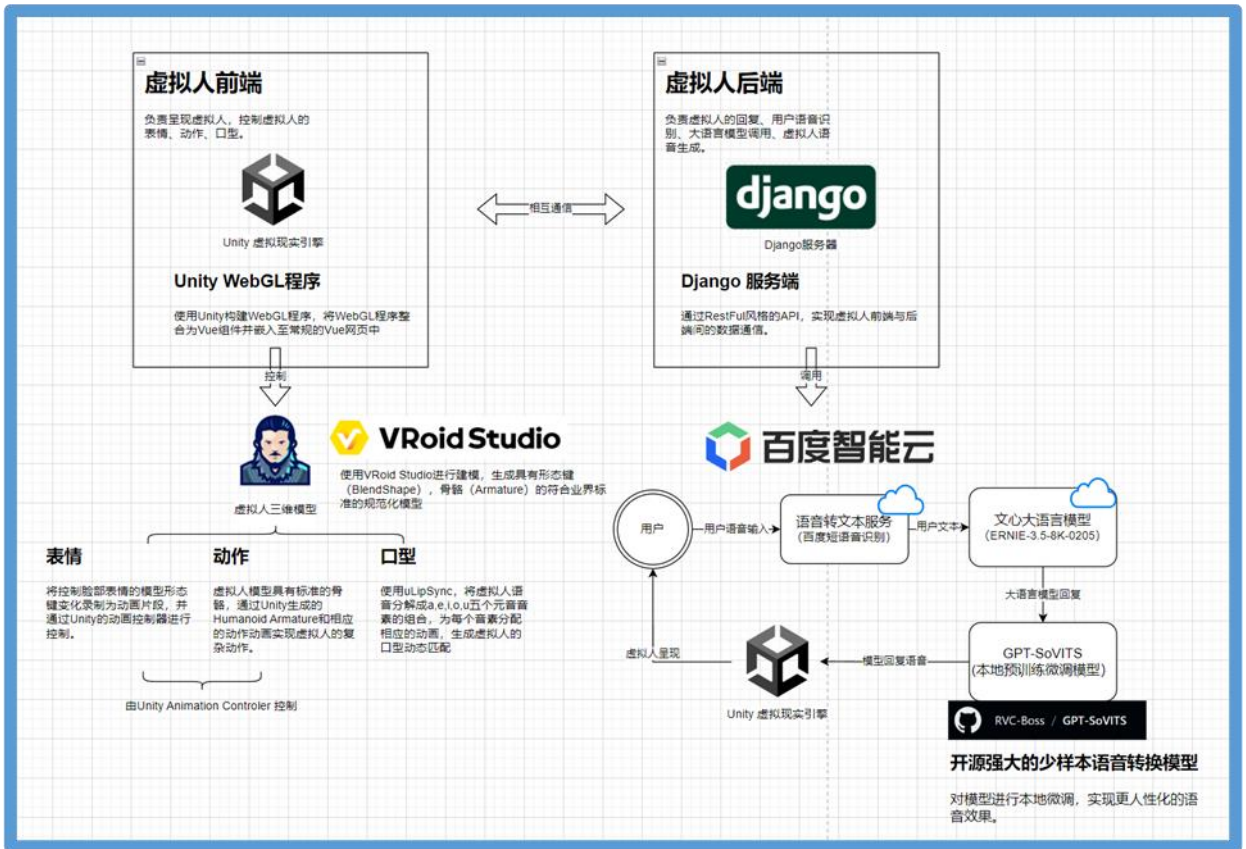


图 6.2.3.1 系统架构图

1、前端

① Unity WebGL 前端

Unity WebGL 前端使用 Unity 虚拟现实引擎开发，其全权负责控制虚拟人的表情、动作、神态。

Unity WebGL 程序负责从 Vue 前端中接受用户的语音或文本输入，随后通过 HTTP 请求把语音输入或文本输入传递给虚拟人后端，虚拟人后端再把回复语音和虚拟人控制信息传输回 Unity WebGL 前端，Unity WebGL 程序从控制信息中提取相应的“表情”、“动作”字段，根据不同的表情和动作，通过虚拟人的 Animation Controller 切换不同的 Animation Clip 从而控制虚拟人的神态改变，在虚拟人的动作改变之间，同时通过调整 Blend In 的相关参数，使虚拟人的动作切换流畅自然。

当 Unity WebGL 程序接收到虚拟人后端传输回的语音信息时，语音将通过虚拟人的 Audio Source 组件发出；同时使用 uLipSync 插件，监听虚拟人的语音，将其语音分解为 A、E、I、O、U 五个基本元音的线性组合，并为每个元音的发声分配口型动画，通过语音与五元音的相似度，动态调整虚拟人的口型，实现虚拟人的口型实时识别效果。

同时 Unity WebGL 程序也会将虚拟人后端发送的控制信息一并传输给 Vue 前端，通过挂载到浏览器 Window 对象上的自定义函数 TransmissionToVue，Unity WebGL 程序将会把控制信息以字符串的形式传输给 Vue 前端，并在 Vue 前端中进行相应处理。

② Vue 前端

Vue 前端使用 Vue3, three.js 开发, 前端的首页是登录界面, 用户可以在此输入账号信息登录账户, 登录界面的左侧是使用 three.js 实现的三维可交互动态模型, 用户可通过鼠标点击进行交互, 增添了网站的趣味性, 登录界面右侧则是常规的账号登录模块, 在登录界面用户即可点击按钮呼唤出虚拟教学助理小慧, 进行实时语音或文字交流。

在用户登录账号后, 即可跳转到网站首页中, 网站首页中间会有大模型生成的每日鼓励致辞, 鼓励用户继续学习。首页下方则是网站提供的精选好课, 用户可点击课程进入课程页面, 查看课程相关信息, 选择课程进行学习等。

网站的顶端是导航栏, 用户可在此切换网站的颜色风格, 网站提供了浅色和深色两种页面风格。用户也可在此进行页面跳转或进入个人主页等。在个人主页页面中用户可以调整自己的账号信息, 修改密码, 个人信息等。

在我的学堂页面中一共有四大模块: “前往学习” 模块, “学习进度” 模块, “做题历史” 模块, “AI 评估” 模块。

在“前往学习”模块中, 会显示平台通过用户画像和用户数据和智能推荐算法推荐的例题, 用户可点击前往做题; 下方则会显示用户所选择的课程, 用户可点选某一门课程进行常规的网络课程学习。

在“学习进度”模块中, 页面将会显示用户各个课程的学习进度, 并以学习圆环的形式动态展示出来; 同时页面也会展示出用户的学习时长情况。并且大模型会根据用户的学习数据, 在每一个图表下给出自己的建议与看法, 以帮助用户更好的把握自身学习情况, 修改学习计划等。

在“做题历史”模块中, 用户可以看到自己以往的做题历史, 并可以选择一道题目进行错题回顾, 重新温习等。

在“AI 评估”模块中, 大模型将会根据用户的所有学习数据和动态学习者画像, 进行综合性评估和反馈, 提出学习建议, 生成个性化学习方案, 供用户加深对自身的学习进度和学习效果的理解, 并进一步精进自己的能力和学识等,

在“题目作答”页面中, 用户根据题目进行作答, 当用户对题目感到困惑迷茫时, 可以点击“提醒”按钮, 让大模型给出题目作答提示, 当用户作答完毕时, 也可点击下方的“评估”按钮, 让大模型根据用户作答情况给出实时评分与反馈, 用户可根据大模型的反馈, 修正作答, 并再次提交进行反馈, 不断循环往复, 直至最终掌握题目。

在“题目作答”页面中还提供了“相似例题生成”功能, 用户可以根据现有题目, 通过本地精调模型生成相似例题, 同时大模型还会一并量化给出题目的相似度, 供用户选择。

2. 后端

平台一共使用了三种后端, 三种后端均使用 Django 实现, 其分别是: “虚拟人服务端”, “数据服务端”, “静态文件服务端”。

系统后端设计如图 6.2.3.2 所示:

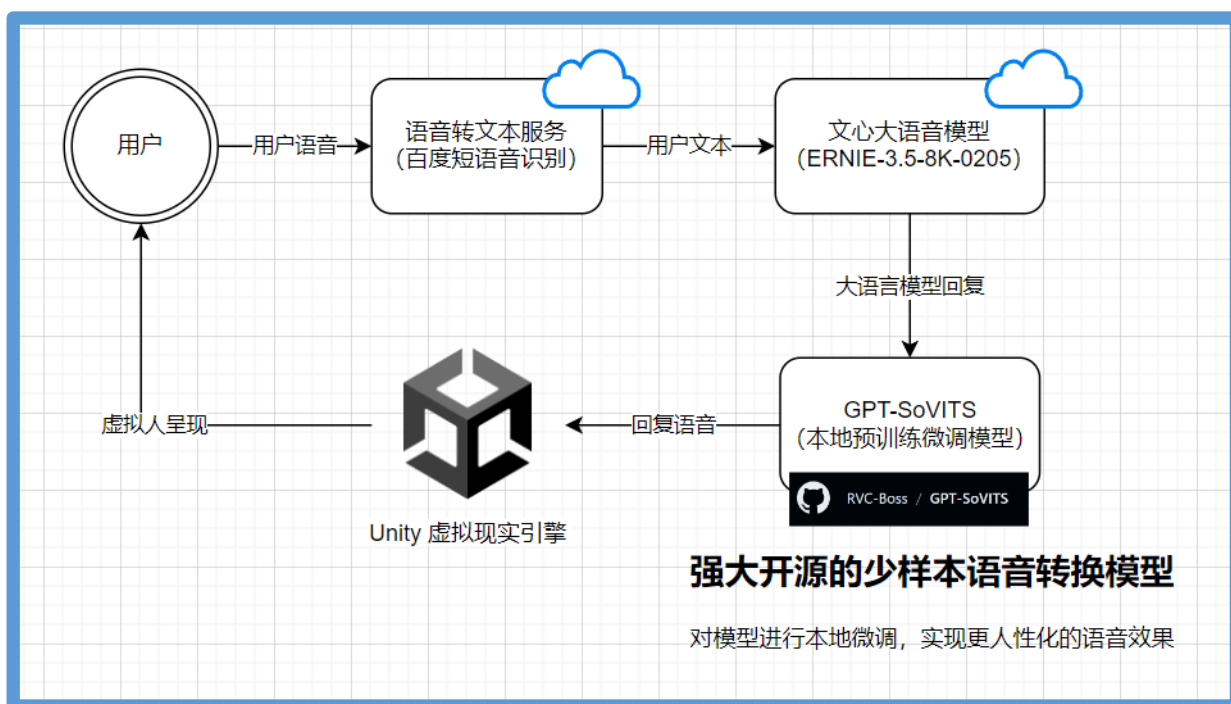


图 6.2.3.2 后端设计图

虚拟人服务端负责控制虚拟人和大模型，主要任务包括：用户语音的识别、大模型回复生成、虚拟人语音合成。在大模型回复生成中，虚拟人后端将根据使用场景的不同，调用不同的大模型进行回复。

当生成相似例题时，虚拟人服务端将调用本地精调的 **Mistral-7B** 生成回复，当用户与虚拟人对话交流、进行数据总结和学习计划生成时，虚拟人后端将调用百度的商用大模型 **ERNIE-3.5** 获得回复，当需要课程知识问答，答疑时虚拟人服务端将调用通过百度千帆大模型工程精调后的 **ERNIE-Speed** 获得回复。而当合成虚拟人回复语音的时候，虚拟人后端将调用本地预训练精调模型 **GPT-Sovits** 进行虚拟人语音合成。

“数据服务端”负责连接 **SQLite** 数据库，提供了常规的网站后台数据管理功能，负责储存网站的课程信息，用户信息，选课信息等，用户的学习数据，选课数据也一并存储在 **SQLite** 数据库中，Vue 前端通过数据服务端暴露出的 **Restful** 风格 **API** 进行通信，对 **SQLite** 数据库进行增删改查。

“静态文件服务端”则是用来储存课程相关静态文件的大容量服务器，其负责存储课程需要的视频、文档、课件等资源，供用户学习，并在需要时提供下载接口等，

七、项目原型

7.1 低保真模型

1. 首页页面



图 7.1.1 首页页面

2. 学习界面

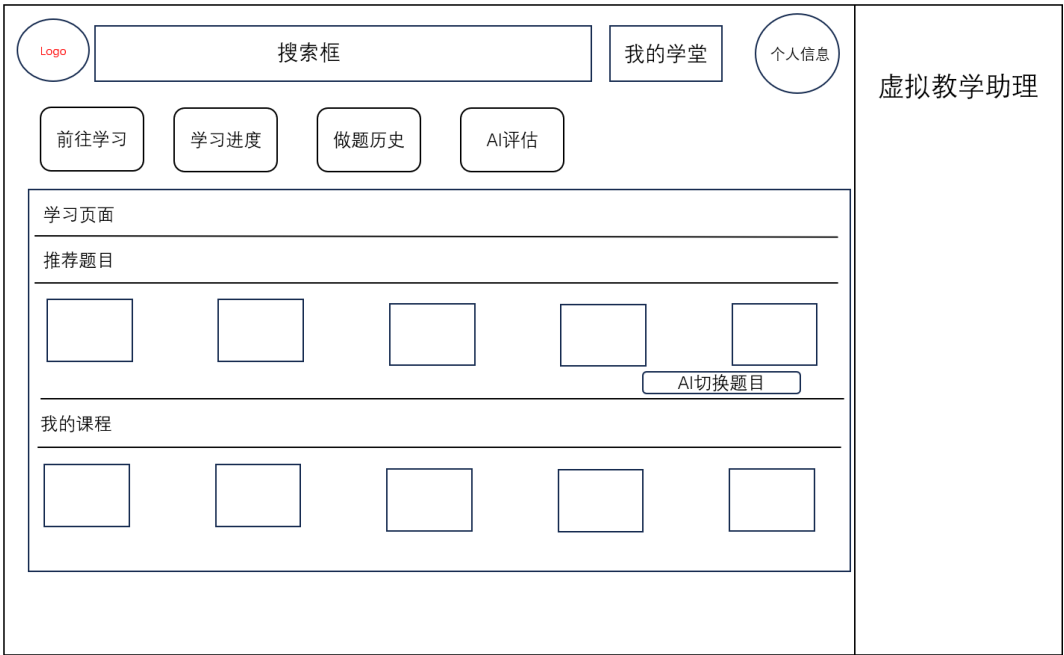


图 7.1.2 学习界面

3. 学习进度展示



图 7.1.3 学习进度展示界面

4. 做题历史界面



图 7.1.4 做题历史界面

5. AI 个性化评估



图 7.1.5 AI 个性化评估

6. 作业详情查看



图 7.1.6 作业详情查看

7. 相似题目生成



图 7.1.7 相似题目生成界面

8. 个人主页

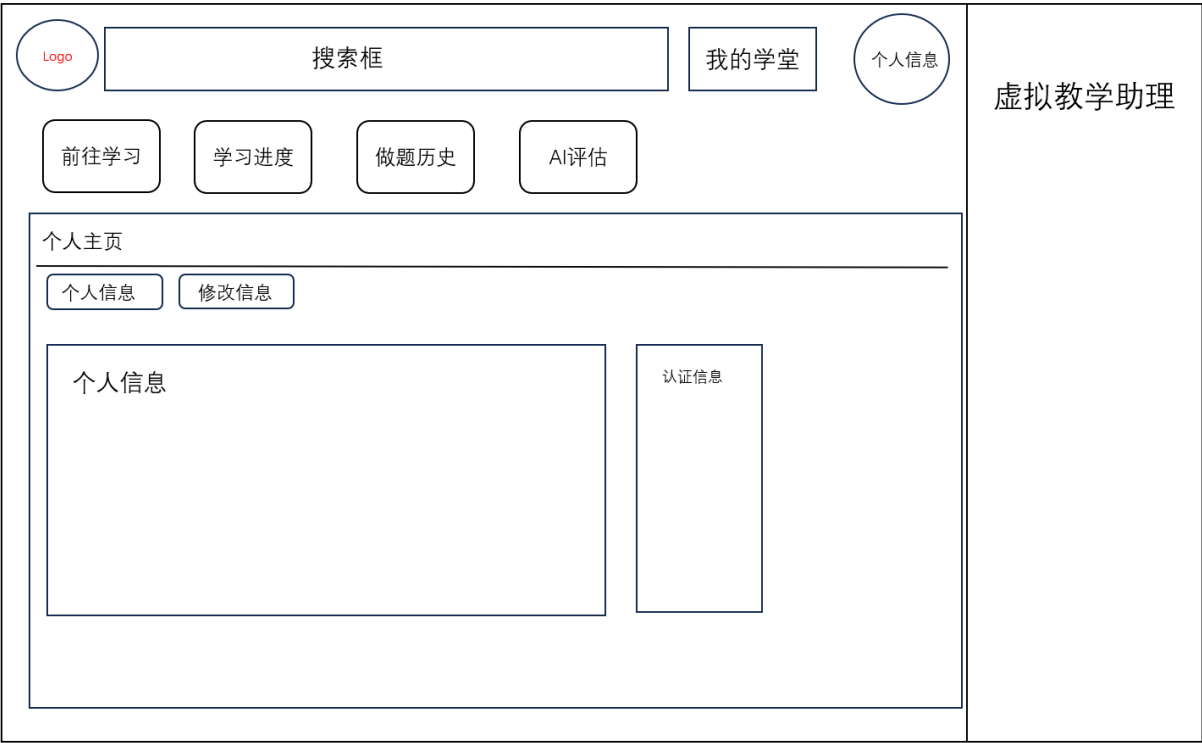


图 7.1.8 个人主页界面

7.2 高保真模型

1. 首页页面

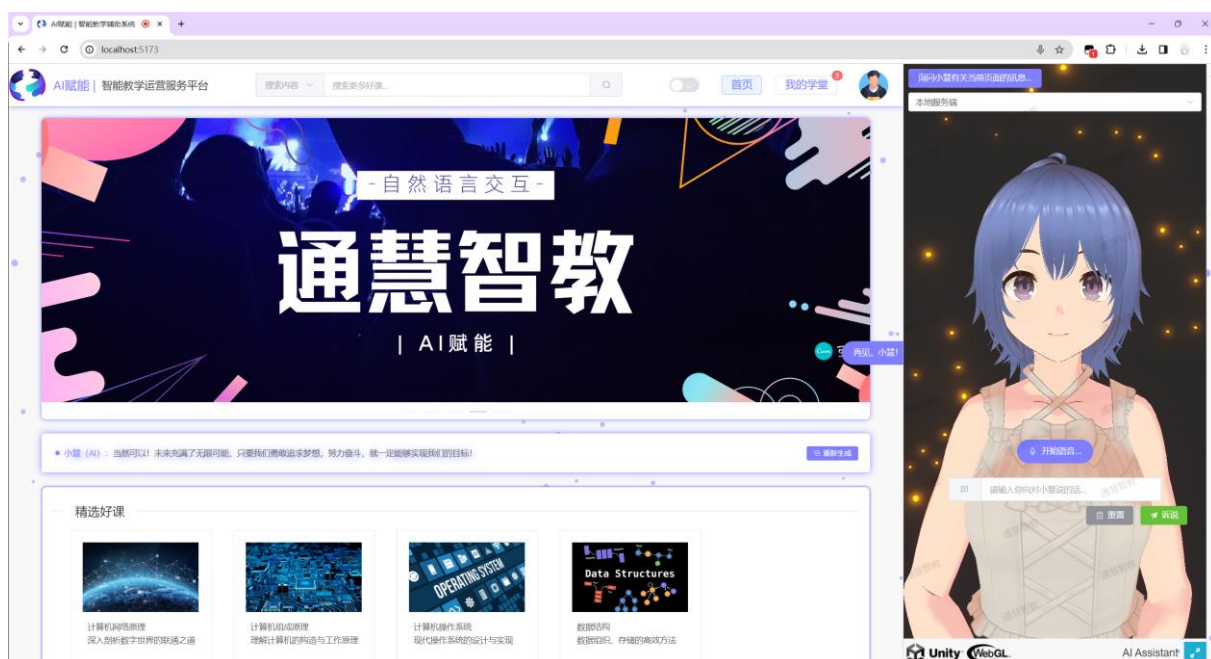


图 7.2.1 首页界面

2. 学习界面

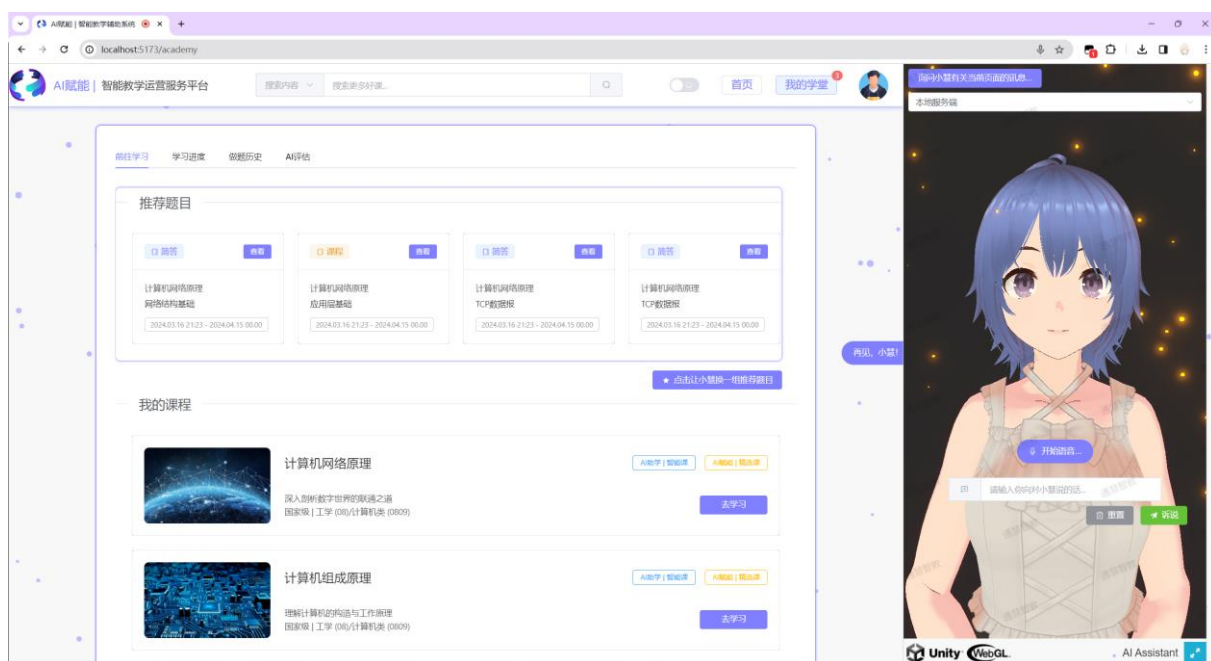


图 7.2.2 学习界面

3. 学习进度展示

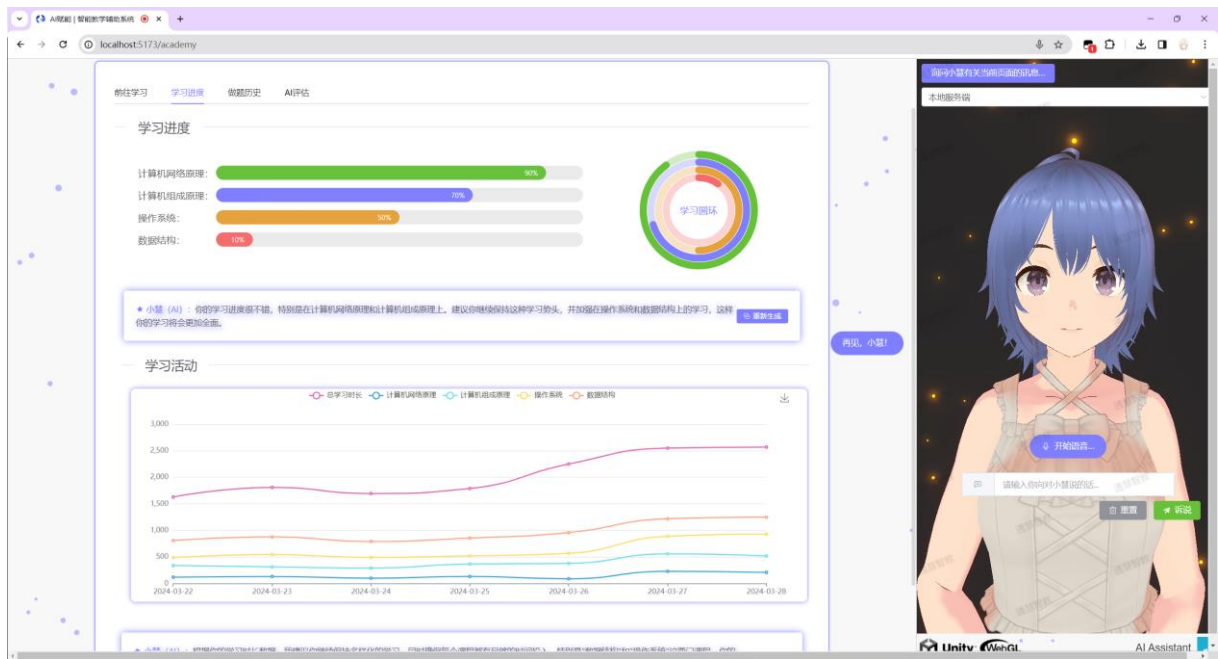


图 7.2.3 学习进度展示界面

4. 做题历史界面

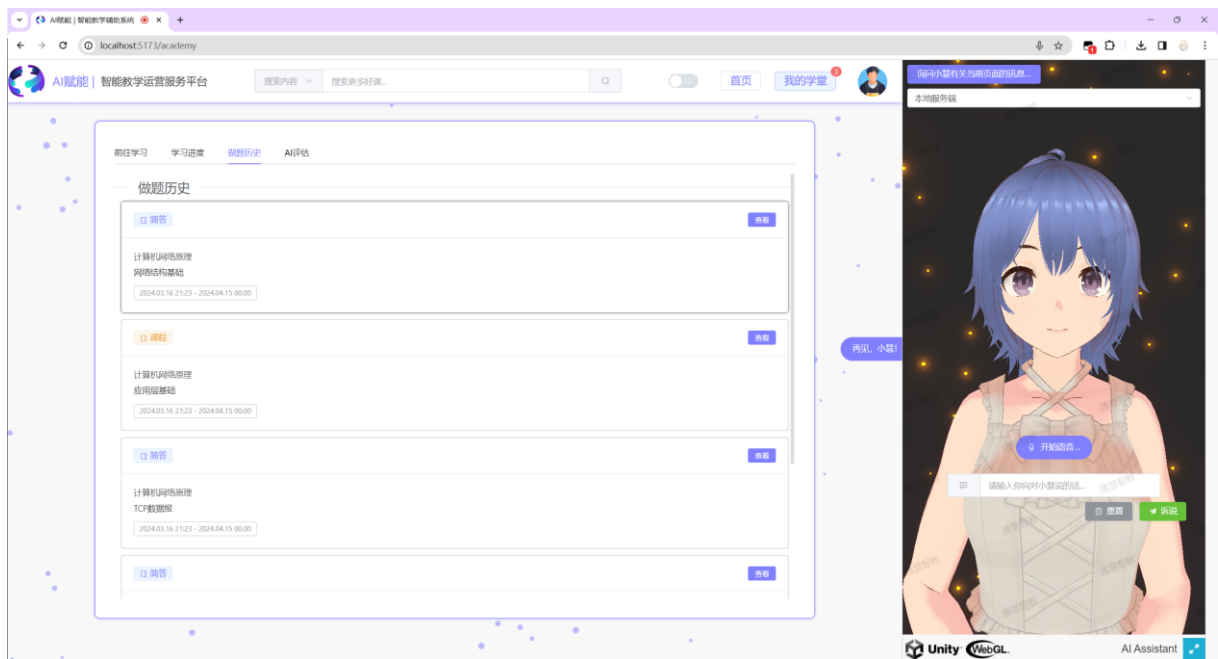


图 7.2.4 做题历史界面

5. AI 个性化评估

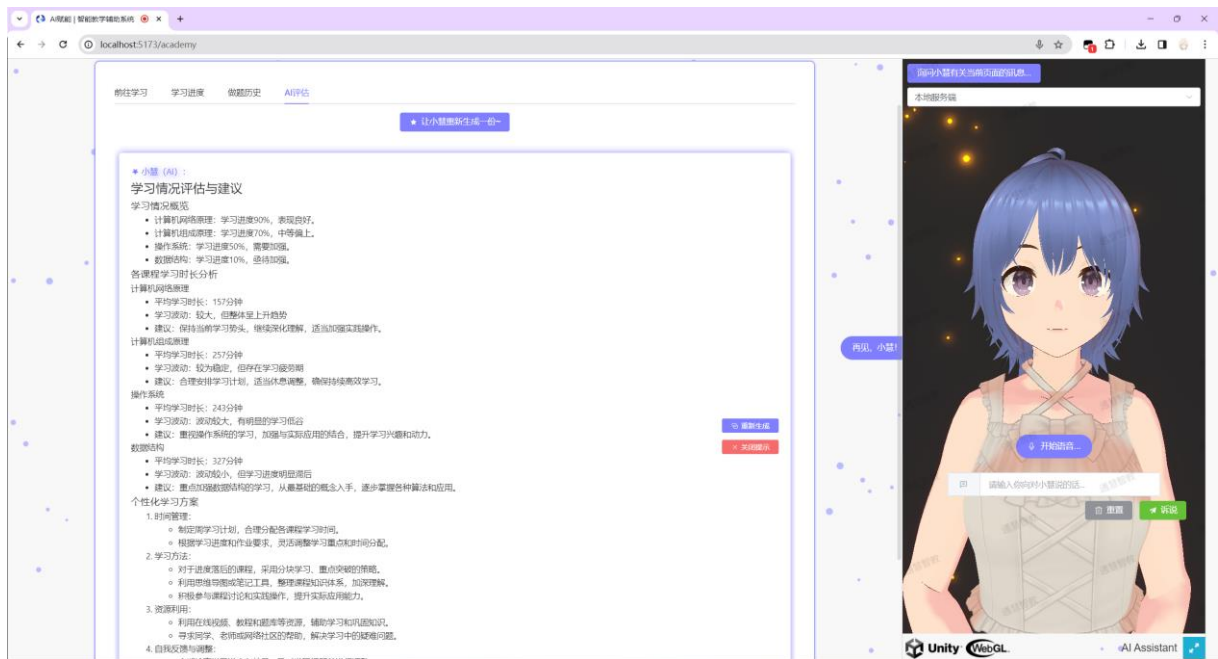


图 7.2.5 AI 个性化评估界面

6. 作业详情查看

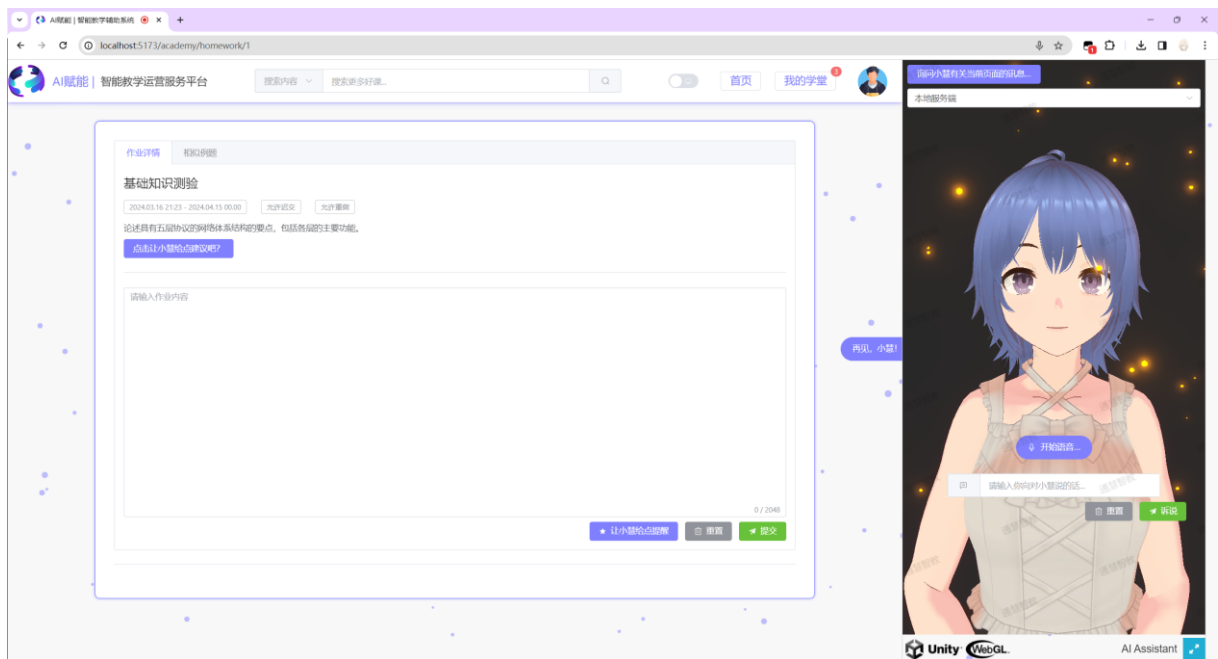


图 7.2.6 作业详情查看界面

7. 相似题目生成

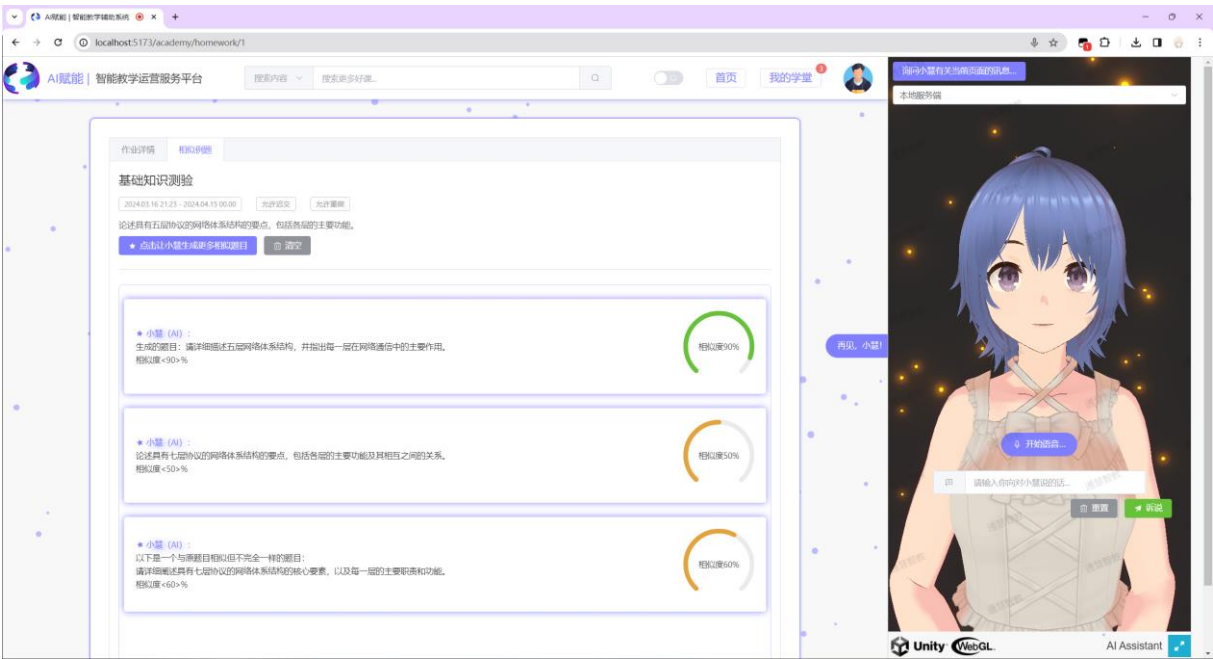


图 7.2.7 相似题目生成界面

8. 个人主页

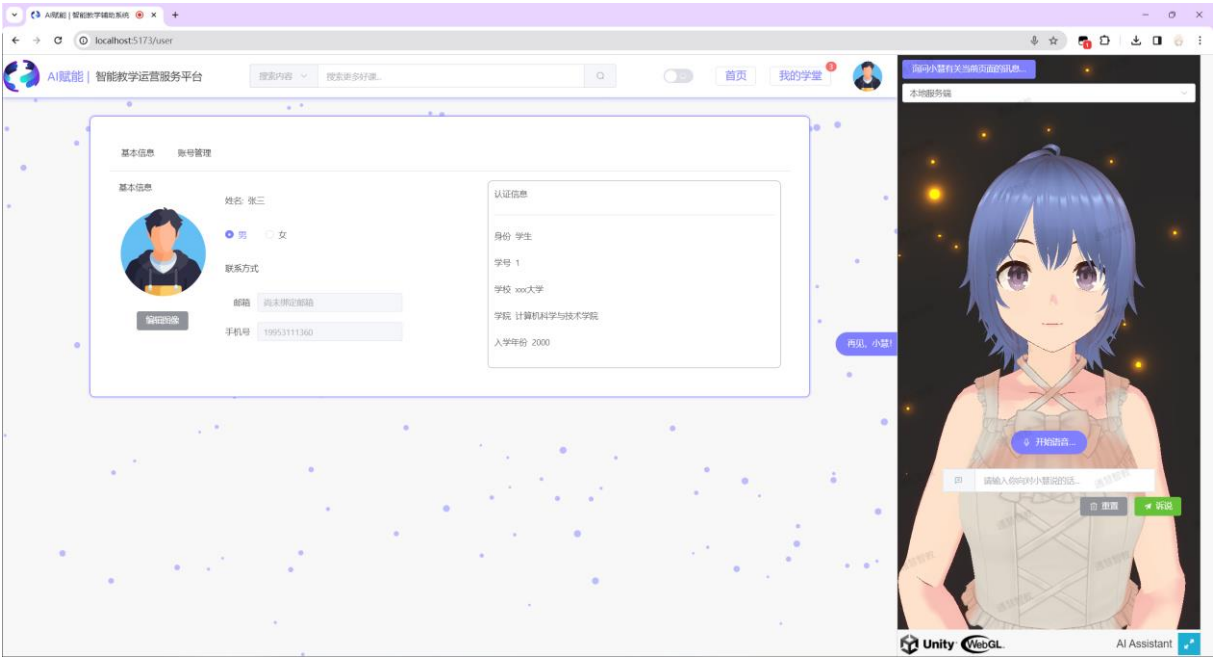


图 7.2.8 个人主页界面

八、开发工具与技术

前端	Vue3、WebGL
算法	LDA、DACF、STT、TTS

后端	Django、Ninja
数据库	MySQL、Redis
建模工具	VRoid、Blender、Unity
虚拟设备	Unity VR、Steam VR
大模型	百度智能云、Mistral-7B

九、应用对象

针对不同的硬件终端，通慧智教系统有不同的应用对象。

9.1 网页端

在通慧智教的网页版平台上，用户可以享受一种全新的学习体验，其中自然语言处理技术使得与虚拟教学助理的对话交流变得流畅自然。利用 AIGC（人工智能生成内容）技术，平台能够根据用户的学习水平和偏好生成个性化的例题，这些例题旨在针对用户的弱点提供专门的练习，从而实现针对性的提升。大模型的实时反馈机制为用户提供即时的学习进度和得分信息，使学习过程更加透明和可量化。

此外，平台还能根据用户的综合学习表现，定制化地生成学习评估报告。这份报告不仅详细记录了用户的学习轨迹，包括所学知识点、练习题表现和学习时长等，还能提供专业的学习建议和改进建议。这种高度个性化的学习体验使得每位用户都能在自己的节奏中高效学习，同时确保了学习内容的相关性和实用性，极大地增强了学习的针对性和有效性。通慧智教平台通过这些先进的技术整合，为用户提供了一个富有创新性和互动性的学习环境。

9.2 VR 端

在 VR 端，用户被引入了一个沉浸式的虚拟环境，其中融合了寓教于乐的小游戏场景，为学习过程带来了趣味性和互动性的提升。这个虚拟空间不仅再现了网页版平台的核心功能，如自然语言交流、个性化例题生成以及实时学习反馈，还通过虚拟现实技术增加了一个维度的体验。用户可以在这个虚拟世界中与教学内容互动，通过参与设计巧妙的小游戏，以游戏的形式掌握知识点，从而达到学习的目的。这种游戏化的学习方法不仅加深了用户对知识的理解和记忆，还大大提高了学习的动机和参与度。定制的学习评估报告同样适用于 VR 环境，为用户提供详细的学习分析和改进建议，帮助用户在享受虚拟现实带来的乐趣的同时，有效地提升学习效果。通过这种创新的学习方式，VR 端为用户打造了一个既丰富多彩又高效实用的学习空间。

十、应用环境

GPU	NVIDIA GeForce GTX 1650 Ti 及以上
CPU	i5-13600KF、R5 7600X 及以上
数据库版本	MySQL5.7、Redis
JDK 版本	1.8
PC 端	Win10 及以上

十一、结语

在面对快速发展的人工智能教育技术（AIGC）所带来的机遇与挑战时，通慧智教通过其创新的方法和措施，展示了如何有效利用 AIGC 技术来推动教育领域的革新。通过精细调整的语言模型、构建高质量数据集、创新教学互动方式以及实现教育内容的个性化，通慧智教不仅提升了教学质量和效果，也为教育资源的优化分配、学习过程的互动性增强、个性化学习的深化发展及学习效果反馈的即时性提供了有效路径。这些成就不仅体现了通慧智教在应对教育领域现有挑战中的前瞻性和创新性，也为未来教育技术的发展和应用提供了宝贵的经验和启示。随着 AIGC 技术的不断进步和深入应用，预期通慧智教将继续引领教育技术革新的潮流，助力构建一个更加高效、个性化和互动性强的教育新生态，激发学生的学习潜力，促进教育公平与质量的全面提升。