

第九届

全国大学生集成电路创新创业大赛

报告类型： 设计报告

参赛杯赛： 算能杯

作品名称： 基于BM1684X的实时课堂思维导图系统

队伍编号： CICC0901957

团队名称： MoE

# 目 录

[第1章 作品介绍 4](#_Toc206624431)

[第2章 系统设计框架 5](#_Toc206624432)

[第3章 系统功能实现 7](#_Toc206624433)

[3.1数据爬取 7](#_Toc206624434)

[3.2 模型微调 8](#_Toc206624435)

[3.3 模型转化与部署 9](#_Toc206624436)

[3.4 教学质量评估 10](#_Toc206624437)

[3.5 网页前端 12](#_Toc206624438)

[3.5 网页后端 12](#_Toc206624439)

[第4章 核心技术和关键指标 13](#_Toc206624440)

[4.1 核心技术 13](#_Toc206624441)

[4.2 关键指标 14](#_Toc206624442)

[第5章 成果 19](#_Toc206624443)

[第6章 创新点 23](#_Toc206624444)

[参考文献 25](#_Toc206624445)

[附录 26](#_Toc206624446)

[1. 生成GNN训练数据使用的fewshot prompt： 26](#_Toc206624447)

|  |  |
| --- | --- |
| 作品名 | 基于 BM1684X的实时课堂思维导图系统 |
| 作品简介 | 本项目基于算能TPU硬件平台，研发了一套面向智慧课堂的实时思维导图生成与智能交互系统。依托搭载BM1684X芯片的Airbox服务器，系统融合语音识别、自然语言处理、大模型语义理解与RAG知识检索技术，实现课堂内容的高效传达与可视化表达。系统采用本地部署与WebSocket低延迟交互，保障数据私密性与稳定性，并由教师端、学生端及本地计算端协同运行。  核心创新包括：①实时将教师授课语音转化为结构化的思维导图，帮助学生捕捉关键信息，并便于课后复习与知识巩固；②基于GNN神经网络理解教师的语义和逻辑关系，开展多维实时评分，解决新教师培训效果难以量化、难以反复练习、评价标准依赖老教师听课等问题；③结合教材知识网络的智慧检索与AI问答，学生可向AI提问，系统自动关联相关知识点并提供思考路径，促进深度理解；④依托端侧算力支持全流程数据本地化处理，避免隐私泄露和网络依赖。  该系统有效提升了课堂教学效率与学生学习体验，同时为教师培训与智能教育提供了安全、高效、可落地的新范式。 |
| 创新点 | 1. **基于课堂语音的实时知识图谱构建技术**   本系统依托Qwen3-4B微调大模型强大的语义理解能力，能够实时解析教师讲课内容中的关键知识点与逻辑关系，自动构建结构化的知识图谱。学生通过图谱可直观掌握知识框架，厘清复杂概念之间的联系，有效提升理解效率。同时，该图谱也为教师提供了教学过程的可视化分析工具，便于动态调整教学策略，实现因材施教。   1. **基于课本知识网络的智慧检索**   系统构建了与教材深度对齐的知识网络，并结合大模型RAG技术，打造了一套智能化的课本级检索机制。该机制不仅支持学生围绕具体知识点进行精准问答，更能从知识体系的角度提供完整、有逻辑的解答路径。与传统检索不同，这种基于知识网络的智能问答为学生提供“解题+溯源”的综合型学习支持，助力深度学习与迁移应用。   1. **基于课堂知识图谱、GNN和统计学方法的课堂质量多维度评估**   系统创新性地构建了多维度、模式自适应的课堂内容思维导图评估标准，涵盖内容完整度、层级清晰度，并针对四种典型图谱模式设定差异化评分区间。通过融合语义与结构特征，实现了高质量数据生成与可量化的智能评估，为教师的授课质量评估提供了科学高效且稳定量化的参考信息。此外，系统还可以使用一些NLP方法和统计学方法分析教师的授课语言习惯并对其给出量化评价。   1. **端侧计算支持下的全流程数据本地化与隐私保护机制**   系统采用端侧计算架构，在无网络或网络受限的环境中亦可稳定运行。所有语音识别、图谱生成、数据分析等关键处理流程均在本地完成，不依赖云端服务器，极大地增强了数据安全性与隐私保护能力。这种全流程本地化部署不仅符合教育场景下对隐私合规的严格要求，也提升了系统的鲁棒性和可达性，为广泛部署提供了技术基础。 |

# 第1章 作品介绍

在当今教育领域，智能教育的蓬勃发展正深刻地改变着传统的教学模式。因此，传统课堂的信息传递方式面临着诸多挑战，一方面在传统课堂中，教师的讲课内容往往存在结构松散、信息密度大等问题，这使得学生在课堂上难以实时捕捉到关键信息，进而影响了他们的学习效果，另一方面对于新教师的培训也存在难以量化标准和难以练习的特点，往往智能通过老教师听课或观摩对比来进行评判。

智能云课堂服务，由于云的特性，隐私数据的安全性难以得到充分保障，存在老师授课信息、学生隐私泄露风险，可能造成不良影响。另一方面，这些服务高度依赖网络，一旦出现断网或网络不稳定的情况，教学活动将无法顺利开展，影响教学效果和学习体验。

为了解决这些亟待解决的问题，本团队借助搭载BM1684X芯片的Airbox服务器成功开发并实现了一套基于算能TPU硬件平台的课堂思维导图实时生成系统。该系统借助先进的技术手段，如实时语音识别、自然语言处理以及可视化展示等，能够将教师的授课内容高效地转化为清晰、直观的思维导图。通过这种方式，系统显著提升了课堂信息的结构化传达能力，帮助学生更高效地获取与理解知识，进行课堂实时提问，从而更好地满足他们在学习过程中的多样化需求。

# 第2章 系统设计框架

该系统由服务器端和客户端（学生端、老师端）两大部分组成，通过 WebSocket 协议实现两端之间的双向通信，从而确保低延迟的实时数据同步能力。这种架构不仅能够高效地处理复杂的教学任务，还为系统的扩展和模块化设计提供了便利，使其能够灵活适应智慧教室、线上直播教学等多种教学场景。

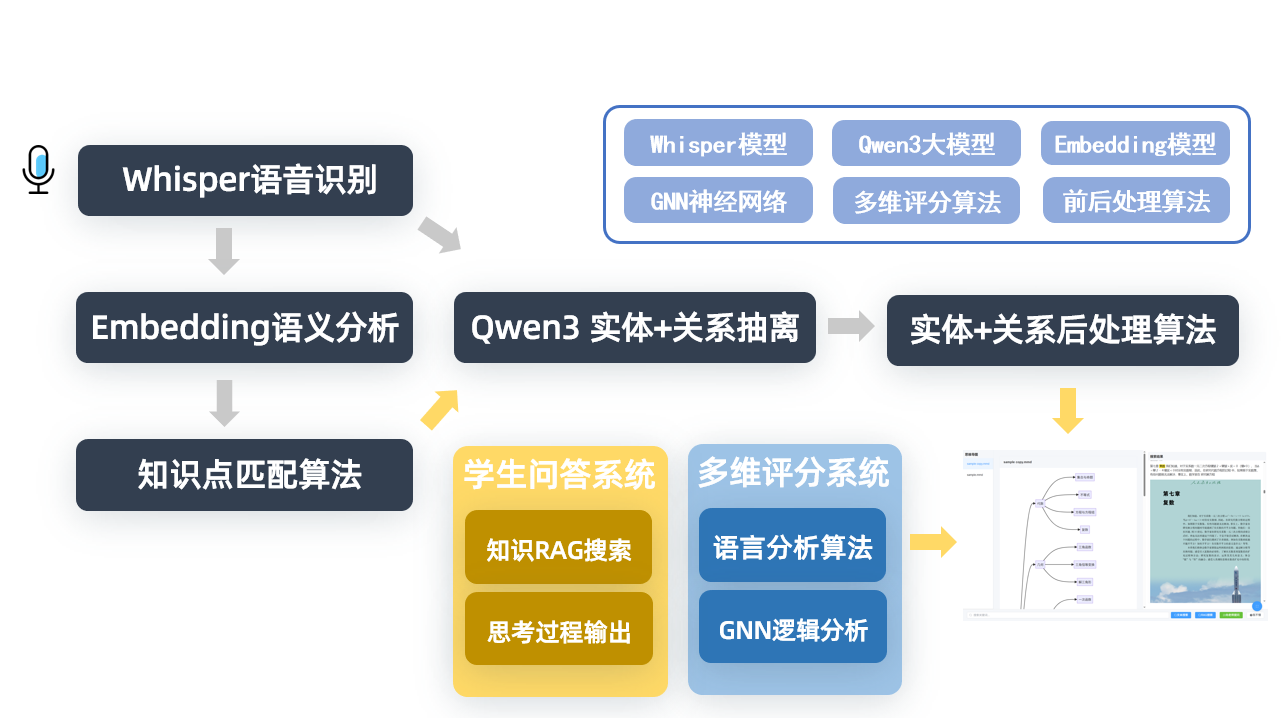


图1-系统架构图

服务器端是系统的核心，承担着音频采集、语音转写、知识点提取、思维导图生成以及 WebSocket 通信等关键功能。其技术实现主要依赖于 Node.js + TypeScript 和 Python 程序的协同工作。Node.js + TypeScript 作为服务器的主框架，负责管理客户端连接、处理 WebSocket 通信、调度任务以及转发数据，同时还提供基础的接口服务和状态管理功能，确保系统的高效运行和稳定性。而Python则专注于音频处理和 AI模型加载任务，通过 pyaudio 实时采集麦克风语音，通过Whisper模型将语音转换为文本，并调用已经转换为bmodel部署在tpu上的经过本团队微调的Qwen3-4B大语言模型解析讲课内容的语义结构，精准提取关键词和逻辑关系，并利用图结构生成标准化的思维导图数据结构Mermaid代码。这些思维导图数据最终通过服务器推送回客户端进行展示。此外，服务器端还集成了基于Qwen3 Embedding的教材索引模块，借助嵌入向量搜索技术，能够根据学生点击的知识点反查课本中相关章节位置，实现语义级别的内容关联，从而为学生提供更加精准的学习资源。我们还使用机器学习技术根据课堂生成的思维导图来为教师的教学提供评价和指导。我们使用大模型对不同的导图依据根据现有相关工作设计的评分标准进行不同角度的打分，并使用这些打分训练了一个计算量更小的GNN网络用于实际使用中对课程进行评价。此外，我们使用NLP方法和统计学方法分析教师的授课语言习惯，从而对授课语言质量（如口头蝉是否过多）进行评价，辅助教师高效快速的改善授课语言质量。

用户端采用 Vue 3 + TypeScript 开发，构建为一个响应式的 Web 应用，具备丰富的功能和良好的用户体验。其核心功能包括实时接收并可视化思维导图、节点交互、教材内容展示以及导图追踪与记录。通过 WebSocket 连接服务器，学生端能够持续获取讲课生成的导图数据，并利用可视化库进行动态渲染，确保学生能够实时查看课堂知识结构。学生可以点击导图中的任意节点，触发客户端与服务器的交互，进一步探索相关知识点。当节点被点击后，客户端会通过 WebSocket 向服务器请求教材内容，服务器则根据教材索引模块返回相关章节，并在客户端下方区域展示，同时支持关键语句高亮、分页跳转等功能，帮助学生更好地理解课堂内容。此外，学生端还支持导图追踪与记录功能，学生可以随时查看过往的导图快照，便于课后复习和摘要整理。在界面设计方面，学生端注重交互友好性和实时性，支持多端访问（如浏览器、平板、白板）。

# 第3章 系统功能实现

## 3.1数据爬取

为了获取微调训练中所需的原始讲课音频数据，我们聚焦于国内最大的视频平台之一——Bilibili。该平台上拥有大量与学习、公开课、知识讲解等相关的优质视频资源。然而，由于其数据体量庞大、结构复杂，传统的单机爬虫方式已难以胜任高效稳定的数据采集任务。因此，我们设计并搭建了一个分布式数据采集系统，以实现对B站海量视频资源的高效抓取、处理与存储。

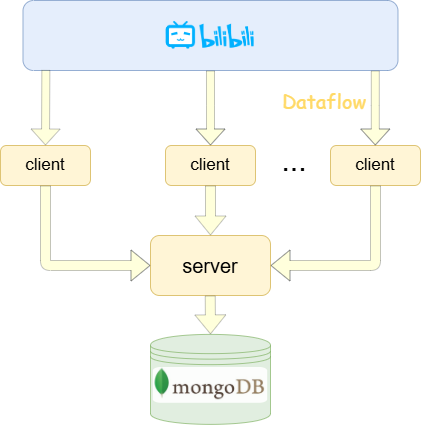


图2-分布式数据采集系统架构图

中心服务器采用Spring Boot框架进行开发，具备良好的模块化结构和高可维护性。其核心职责包括视频信息检索，通过访问B站开放接口（或中转API），获取目标视频的基础信息，如 BV号、标题、UP主、播放时长等；同时负责任务生成与调度，将上述信息打包为结构化任务（例如按频道、关键词、日期筛选），并分发给客户端爬虫节点；此外，服务器还承担任务状态监控，追踪每个任务的处理状态，确保数据完整性与系统健壮性；最后，在客户端上传采集数据后，服务器会统一完成验证、去重及持久化存储。我们选用了 MongoDB 作为后端数据库，其天然支持文档型存储，结构灵活，特别适合处理视频元数据、弹幕、字幕等非结构化信息，同时在高并发写入场景下表现优异，极大提升了采集系统的数据吞吐能力。

客户端模块使用Python 编写，借助其优秀的网络库和多线程支持，能快速搭建高效、易部署的爬虫程序。每个客户端都周期性地向中心服务器请求任务，并独立完成视频页面访问与音频提取，自动访问目标视频页面，使用工具如ffmpeg提取音频流；同时进行字幕与弹幕数据解析，调用接口获取弹幕/字幕，将ASS/SRT格式转换为JSON格式弹幕；接着进行数据清洗与结构化处理，将获取到的信息规范化为统一格式，方便后续存储与训练；最后将音频、字幕等数据打包后上传至中心服务器。为提升效率，客户端内部实现了多线程任务队列与失败重试机制，并具备一定的本地缓存与断点续传能力。

在分布式架构的支持下，系统能够同时调度数十个爬虫客户端，协同完成大规模视频采集任务。在上述系统下，我们高效采集超过10万条有效字幕数据，其中包括课堂讲解、教育访谈、学术分享等内容。在性能方面，我们针对任务进行了专项优化，对请求限速做了防护；同时实现了任务分发负载均衡，基于客户端状态与速率动态调整任务下发密度。

## 3.2 模型微调

为了能够加强大模型对特定任务（生成结构性思维导图mermaid代码）的处理能力，本团队对Qwen-4B模型进行微调。由于终端算力和显存有限，为了让其上运行的小模型对于我们特定的下游任务有更强的效用，我们在更大的模型上生成了一些数据，清洗数据确保生成的mermaid代码没有语法错误并使用这些数据微调小模型，从而实现蒸馏更大模型的文本总结与格式化输出能力。

我们选用的教师模型是Qwen3-32B稠密模型。由于文本总结并生成mermaid对模型的推理能力和思维能力要求不高，且由于deepseek-V3规模过大只能调用API生成蒸馏数据，考虑到成本我们认为Qwen3-32B是较为合理的选择。我们使用如下自定义模版使用Qwen-32B-int8生成了3.7万条数据用于蒸馏：

<|im\_start|>system\n你是一个Mermaid代码生成专家，用户将输入一段录音的语音识别结果，1用中文，但语法关键字（如graph TD、-->）保持英文；\n3. 严格遵循Mermaid语法规范。\n\n注意语法细节：\n- 节点定义格式：节点名[中文标签]，如 A[用户登录]\n- 连接符：使用 -->、---、==>|文字|== 等标准符号\n- 子图需用subgraph包裹并正确闭合\n- 方括号中的文字（即节点标签）内容请用引号包裹\n- 节点名称请使用不会与mermaid语法发生冲突的无意义内容，节点需要显示的内容请放在节点标签（即名称后的中括号）中\n\n\*请确保你生成的是Mermaid代码!你只需要生成Mermaid代码，请不要附加其他信息!Mermaid代码请放到代码块中!\*\n\n语音识别结果如下：\n<|im\_end|>\n<|im\_start|>user\n

[语音识别文本]

\n<|im\_end|>\n<|im\_start|>system\n现在请你将上述内容详细的总结并绘制一张思维导图，用Mermaid形式，生成符合语法的Mermaid代码。要求：\n1. 仅输出标准Mermaid代码，不包含任何解释或注释；\n2. 节点名称需用中文，但语法关键字（如graph TD、-->）保持英文；\n3. 严格遵循Mermaid语法规范。\n\n注意语法细节：\n- 节点定义格式：节点名[\"中文标签\"]，如 A[\"用户登录\"]\n- 连接符：使用 -->、---、==>|文字|== 等标准符号\n- 子图需用subgraph包裹并正确闭合\n- 方括号中的文字（即节点标签）请用引号包裹\n- 节点名称请使用不会与mermaid语法发生冲突的无意义内容，节点需要显示的内容请放在节点标签（即名称后的中括号）中\n\n\*请确保你生成的是Mermaid代码!你只需要生成Mermaid代码，请不要附加其他信息!Mermaid代码请放到代码块中!节点标签中的文字请使用英文引号\"\"包裹!\*<|im\_end|>\n<|im\_start|>assistant\n<think>\n\n</think>\n\n

图3-模型微调自定义模版

为了提升蒸馏数据的生成速度，考虑到文本总结和mermaid图象绘制任务中引入reasoning thinking可能会导致输出不稳定，我们生成数据时没有使用reasoning thinking。为了提升数据质量，我们去除了过长（大于9500个token）和过短（小于150个token）的数据。

生成数据后，为了防止小模型学习到噪声，我们使用mmdc校验每条生成的数据是否有mermaid语法错误，并去除了其中存在语法错误的数据。筛选过后，还剩余3.6万条数据。

我们选取了两款蒸馏的目标模型，分别是是Qwen3-4B和Qwen3-8B。该模型可以在tpu上高速推理。我们使用QLoRA以适应最终的量化运行环境，并节省微调使用的计算资源。QLoRA使用8bit bnb量化，lora\_rank=32, lora\_alpha=16, lora\_dropout=0.1, LR=5e-4, epoch=2, 并使用梯度积累实现等效batch\_size=16。微调之后，我们将LoRA合并到base model中并导出safetensors模型。

## 3.3 模型转化与部署

在获得模型后，进行模型的转换工作。使用tpu-mlir工具链，在x86工具站拉取docker镜像后使用工具llm\_convert.py进行转换，使用llm\_convert.py -m /workspace/Qwen3-4B -s 512 --quantize w4bf16 -c bm1684x --out\_dir qwen3\_4b命令将HuggingFace的safetensor格式的模型转换为为 Top 层(硬件无关层)mlir 模型，接着将mlir模型转换为可在tpu进行部署工作bmodel格式，使得模型能够在服务器tpu上正常运行。模型转换流程示意图如下图所示：

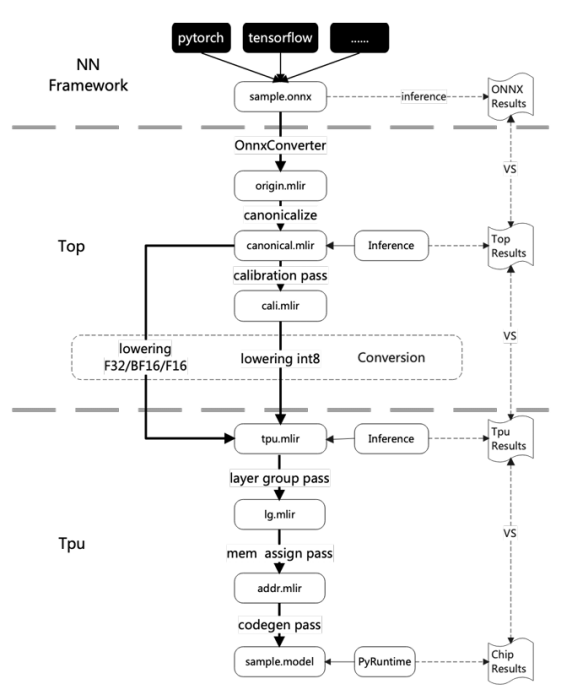


图4-模型转换结构示意图

获得微调的模型后，参考示例工程代码编写部署代码，使得微调后的Qwen3-4B模型能够正常，从而能够正常运行，将Whisper通过语音转录得到的txt文件转换为思维导图。

## 3.4 教学质量评估

我们的工作允许程序对根据思维导图数据对教师的课堂质量进行打分并提供修改建议，从而帮助教师提升教学质量。为了降低硬件开销，我们将评估模型设置为对思维导图进行打分的GNN。

为确保思维导图评分的客观性与一致性，参考[5-9]我们建立了一套多维度的综合评估体系。该体系围绕四大核心维度展开：结构逻辑性 (35分)，用于评估图谱的框架合理性与知识关联度；内容完整度 (35分)，旨在衡量知识点的深度、准确性与精炼程度；层级清晰度 (20分)，关注图谱的视觉呈现与层次关系；以及代码规范性 (10分)，评估Mermaid代码本身的质量。通过为“详细型”、“逐渐展开型”、“简略大纲型”和“直线型”这四种典型模式设定明确的分数区间，该标准不仅能量化评估每一份作品的质量，更能有效地区分不同复杂程度的思维导图，为后续的机器学习模型训练提供了坚实、可靠的数据基础。

我们使用大模型依据上文设计的评估体系和我们的应用场景生成了大量数据，并对其生成的思维导图数据进行打分，从而得到GNN训练数据。整个流程首先从九大学科（生物、数学、语文等）中精心挑选出数百个核心知识点。随后，我们利用大型语言模型，针对每一个知识点，系统性地生成了四种不同类型的Mermaid思维导图：结构严谨、内容详尽的“详细型”；层次清晰、逐步深入的“逐渐展开型”；框架清晰、要点明确的“简略大纲型”；以及逻辑线性、步骤简单的“直线型”。每一份生成的导图都由模型依据预设的评分标准进行自我评估打分，确保了标签的一致性。最终，我们通过对现有知识点的循环利用和引入新的通用知识点，成功生成了数千条高质量、带有多维度评分标签的思维导图数据，为模型训练提供了充足的“养料”。为了提升数据的多样性从而提升模型的稳健性，我们也采用了一部分蒸馏Qwen3-4b时使用的数据并使用LLM根据我们的评分标准进行了打分放入训练数据中。

为了让图神经网络（GNN）能够同时理解思维导图的语义内涵与拓扑结构，我们设计了一种创新的多模态特征融合策略。首先，针对节点特征，我们将通过SentenceTransformer模型生成的384维高维文本嵌入，利用主成分分析（PCA）技术降维至16维。这一步骤旨在保留核心语义信息的同时，显著降低文本特征的维度，避免其在训练中压制结构信号。随后，我们为每个节点计算了5个关键的局部结构特征（如度、中心性等），并与16维的文本特征拼接，构成最终的21维节点输入。在图级别，我们进一步提取了9个描述思维导图宏观形态的全局图特征（如密度、直径、线性度等），并进行归一化处理。在模型架构上，我们采用图注意力网络（GAT）来学习节点间的复杂关系，并通过Set2Set池化层聚合节点信息生成图嵌入。最终，池化后的图嵌入将与9个全局图特征拼接，共同送入一个MLP决策层，从而实现对思维导图质量的精准、全面的评估。

由于我们的GNN模型结构简单且调用次数显著少于本地LLM，我们直接使用CPU进行计算。LLM生成数据时使用的fewshot prompt见附录

本项目还提供了课堂录音语言习惯分析功能，用于改善教师授课的语言习惯问题。系统使用NLP算法和统计方法统计教师的口头蝉并对录音转文字的文本进行全面评估，从而可以给出教师授课过程中的不佳语言习惯，辅助教师快速改善授课语言质量。

## 3.5 网页前端

客户前端采用 Vue 3 + TypeScript 开发，构建为一个响应式的 Web 应用，具备良好的用户体验和多端访问能力（如浏览器、平板、白板）。其核心功能包括实时接收并可视化思维导图、节点交互、教材内容展示以及导图追踪与记录。学生端通过 WebSocket 持续获取服务器推送的导图数据，并利用可视化库动态渲染，支持点击节点触发交互、教材内容高亮与跳转等功能，便于学生实时查看课堂知识结构和课后复习。

## 3.5 网页后端

服务端主要基于 Node.js 和 Express 框架开发，主要职能包括静态资源托管、WebSocket 实时通信、PDF 检索以及 Mermaid 文件管理等。服务端通过 Express 提供静态资源服务，向客户端分发前端页面、PDF 图片、Mermaid 文件等，其中 Mermaid 文件的增删改由 fileWatcher.ts 负责，利用 chokidar 监听文件系统变化，实时推送文件列表和内容变更给客户端，以保证前后端数据一致。此外，服务端集成了 socket.io，实现与客户端的双向实时通信，所有与文件、检索、互动相关的事件（如文件列表更新、PDF 检索请求、学生提问、老师广播等）都通过 WebSocket进行，事件处理逻辑集中在socketHandlers.ts，根据不同事件类型分发处理函数，确保消息能够及时、准确地传递到对应客户端。PDF检索功能由pdfSearcher.ts实现，服务端预先将PDF教材按页切割为图片，并建立关键词索引，收到检索请求后，会在索引中查找相关内容，返回匹配的页码、图片路径和上下文摘要，便于前端高亮展示和定位。服务端还负责收集和统计学生的提问、“我不懂”等互动数据，并实时推送给老师端，所有互动数据通过 WebSocket 事件传递，支持多用户并发和状态同步。服务端支持 Docker 容器化部署，方便在不同环境下快速部署和扩展。

通过上述设计，系统实现了师生间的高效知识检索与互动，支持多用户并发、实时数据同步和灵活的知识图谱展示，极大地提升了教学与学习体验。

# 第4章 核心技术和关键指标

## 4.1 核心技术

系统的全部组件均在BM1684X上运行，主要有以下几部分：Whisper实时语音转写、Qwen3实时思维导图建图、前后端思维导图渲染、老师学生实时互动、课堂质量评价。

Whisper模型：Whisper模型负责将教师的实时语音输入快速转换为文本内容。它基于先进的语音识别技术，能够高效且准确地将语音数据转化为文字，为后续的思维导图建图和知识结构化提供基础数据支持。Whisper 模型在 BM1684X 上运行，确保了低延迟和高效率的转写体验，让教师的每一句话都能被及时记录并转化为可处理的文本信息。

Qwen3实时思维导图建图：Qwen3 实时思维导图建图功能利用微调过的 Qwen3-4B 大模型的强大理解能力，在 BM1684X 本地进行推理。该功能能够自动解析教学内容的语义结构，提取关键知识点和逻辑关系，并生成清晰的知识图谱。这种图谱不仅帮助学生更直观地理解复杂知识，还能为教师提供教学分析工具，助力优化教学策略。随着教师讲解的进行，Qwen3 模型能够实时生成思维导图，动态展示知识的层次和关联，便于学生实时跟进和理解。同时，生成的知识图谱可以作为教师的教学分析工具，帮助教师评估教学内容的逻辑性和完整性，从而优化教学策略，提升教学质量。

前后端思维导图渲染：本系统采用高效的前后端协同渲染技术，实现思维导图的动态生成与实时展示。前端基于 Vue 3 + TypeScript 构建响应式 Web 应用，利用 Mermaid.js 等可视化库，将服务器推送的思维导图数据实时解析并渲染为直观的图形结构，支持多端访问与动态更新。后端则通过 Node.js + Express 框架，结合 WebSocket 协议，确保导图数据的低延迟传输与实时同步，保障学生端能够无缝接收并展示最新导图内容。同时，系统支持节点交互功能，学生点击任意节点即可触发教材关联展示与进一步知识探索，极大地提升了知识传达效率与学习体验。

老师学生实时互动：本系统通过 WebSocket 协议实现低延迟的双向通信，为师生提供高效的实时互动功能。学生端支持在思维导图节点上直接点击提问，问题通过 WebSocket 即时传输至老师端，老师可实时查看并针对性解答。同时，学生还可点击“我不懂”按钮，系统实时统计并向老师反馈不懂数量，帮助老师精准把握教学节奏，调整授课内容。此外，系统支持文本向量搜索与 RAG 搜索功能，学生可在客户端输入问题，系统快速检索相关知识点并展示，辅助学生自主学习。这种实时互动机制极大地提升了课堂教学的互动性和教学效果，满足了智慧教室场景下的多样化教学需求。

课堂质量评价：本系统通过预先训练的GNN模型对整个课堂不同方面的质量进行打分评价。可以辅助教师检查课堂的逻辑是否清晰明了、内容是否充实、知识讲解层级是否清晰，结合学生疑问等数据辅助教师改进课堂质量。本系统还使用了一些NLP和统计方法统计教师的语言习惯，从而可以辅助教师改善语言质量，提升课堂质量。

## 4.2 关键指标

Qwen3-4B模型微调的loss曲线如下：

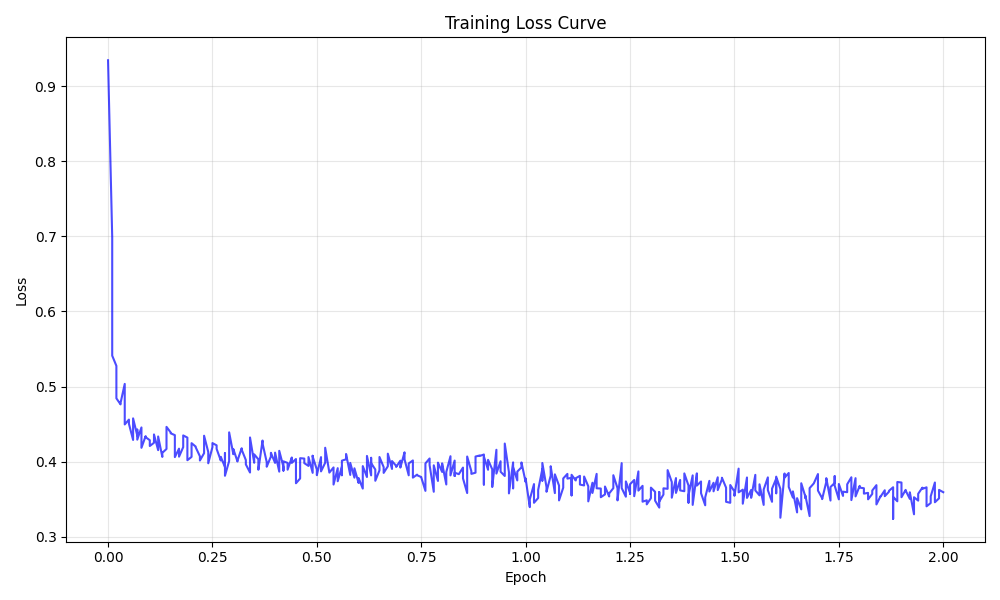


图5-模型训练loss曲线图

从loss曲线可知模型已经基本收敛。模型微调后测试集上的loss为0.39，说明4B的小模型已经充分从32B的教师模型处习得相关知识。经过实际测试，Qwen3-4B模型可以正常的输出高质量且无语法错误的mermaid导图代码，而微调前容易出现语法错误或总结不全面的情况。

下面对微调模型与其他模型进行对比，说明模型微调的效果。使用一样的prompt对不同的模型进行输入，得到的结果如下：

Deepseek-671B模型输出效果如下：

graph TD A["细胞分裂专题"] --> B["细胞分裂的原因"] A --> C["染色体与基因的关系"] B --> D["相对表面积减小"] B --> E["核质不均衡"] B --> F["DNA复制"] D["相对表面积减小"] --> G["细胞过大导致物质运输困难"] G --> H["必须通过分裂解决"] E["核质不均衡"] --> I["细胞核无法控制整个细胞质"] I --> J["遗传信息不足以支持细胞需求"] J --> K["细胞分裂是唯一解决方案"] F["DNA复制"] --> L["遗传物质需要复制后才能分配"] L --> M["否则染色体数量会减少"] M --> N["可能导致变异，如21三体综合征"] C["染色体与基因的关系"] --> O["性状的定义"] O --> P["性状是由基因控制的"] P --> Q["一个性状可能由多个基因共同控制"] C --> R["基因与DNA的关系"] R --> S["基因是有遗传效应的DNA片段"] S --> T["每个DNA上包含多个基因"] C --> U["染色体组的概念"] U --> V["一个染色体组 = 含有本物种一整套遗传信息的一组染色体"] V --> W["不同染色体功能不同"] W --> X["一条染色体负责头部发育"] X --> Y["另一条染色体负责颈部发育"] Y --> Z["依此类推，组成完整的身体结构"] U --> AA["染色体组的特征"] AA --> AB["大小、形态、结构、功能各不相同"] AB --> AC["每条染色体携带不同的基因集合"] AC --> AD["共同控制生物的全部性状"]

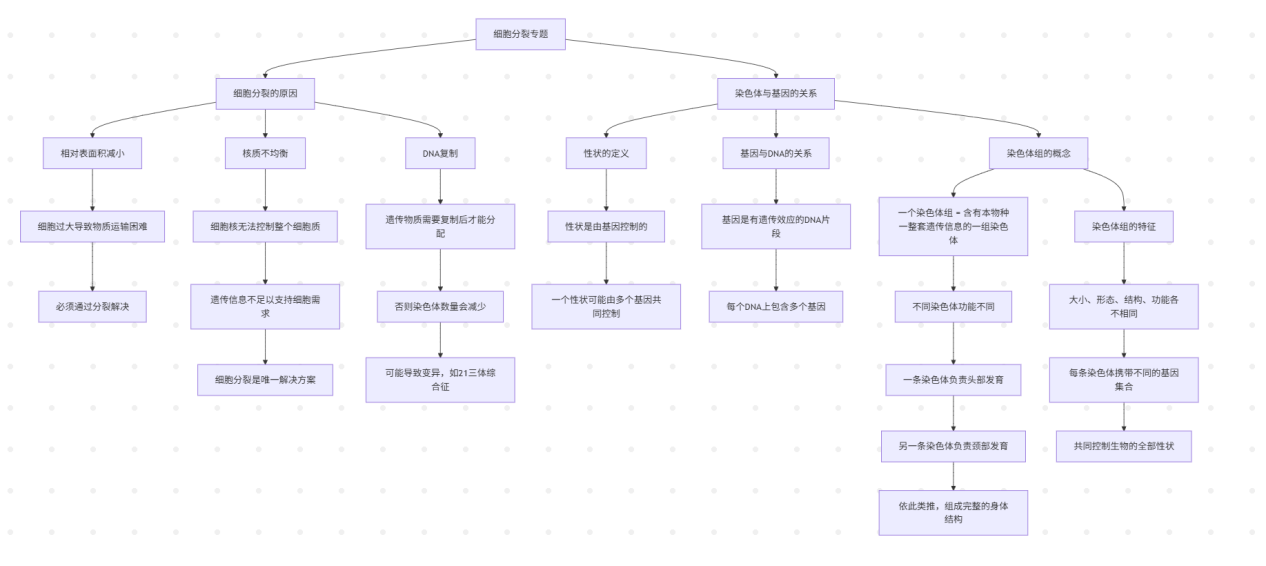


图6-deepseek-671B运行结果图

Qwen-32B模型输出效果如下：

graph TD A["细胞分裂的原因"] --> B["相对表面积减小"] A --> C["核质不均衡"] A --> D["DNA复制"] B --> E["细胞体积增大导致相对表面积减小"] E --> F["物质运输效率降低"] C --> G["细胞核对细胞质的控制能力不足"] G --> H["遗传物质调控蛋白质合成的需求增加"] H --> I["蛋白质需求超过合成极限"] D --> J["确保遗传物质平均分配"] J --> K["染色体复制形成姐妹染色单体"] J --> L["避免遗传物质分配异常(如21三体综合征)"] A --> M["性状与基因"] M --> N["性状由基因控制"] N --> O["基因是具有遗传效应的DNA片段"] N --> P["基因数量远多于性状数量"] O --> Q["例:玉米性别由6对基因控制"] M --> R["DNA分装机制"] R --> S["长DNA链分装为多条染色体"] S --> T["染色体组定义"] T --> U["包含物种全套遗传信息"] T --> V["染色体形态功能各异"] T --> W["例:人类染色体组控制头/颈/四肢发育"]

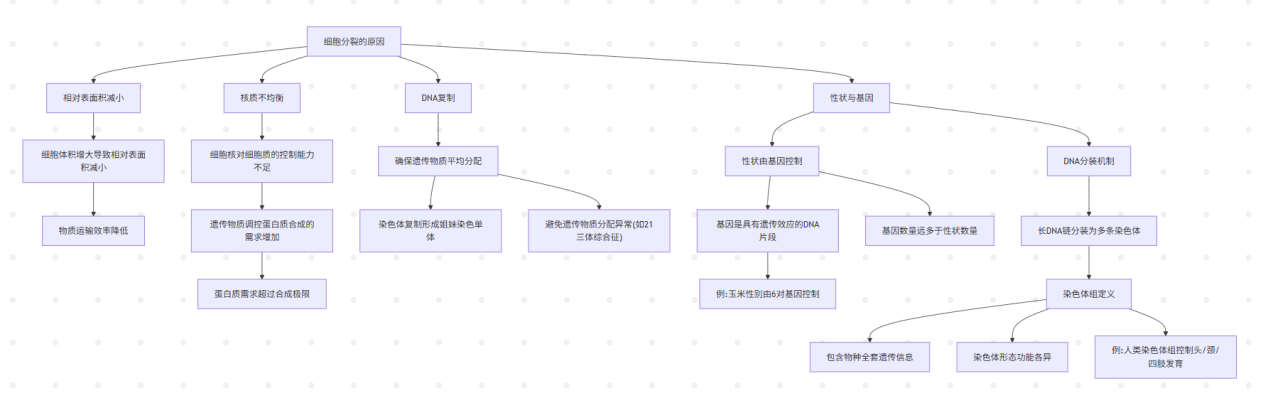


图7-Qwen-32B模型运行结果图

Qwen-4B模型输出效果如下：

graph TD A["细胞分裂专题"] A --> B["细胞分裂原因"] A --> C["细胞分裂现象与性状"] B --> B1["相对表面积减小"] B --> B2["遗传物质不均衡"] B --> B3["DNA复制"] B1 --> B1a["细胞体积增大导致表面积/体积比下降"] B1 --> B1b["物质运输效率降低"] B1 --> B1c["细胞分裂必要性"] B2 --> B2a["细胞核控制能力有限"] B2 --> B2b["遗传物质分配不均"] B2 --> B2c["细胞分裂解决矛盾"] B3 --> B3a["DNA复制保证遗传物质均等分配"] B3 --> B3b["分裂前必须完成复制"] B3 --> B3c["复制失败导致变异风险"] C --> C1["性状与基因关系"] C --> C2["染色体组概念"] C1 --> C1a["性状由基因控制"] C1 --> C1b["多性状由多基因控制"] C1 --> C1c["基因数量与性状数量差异"] C2 --> C2a["染色体组定义"] C2 --> C2b["染色体组功能差异"] C2 --> C2c["染色体组组成原理"]

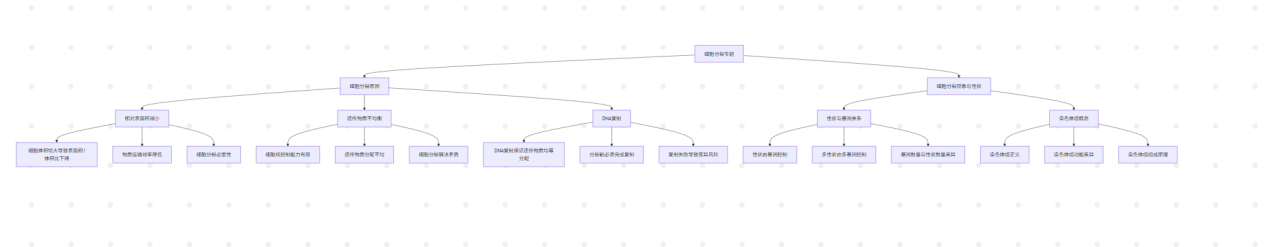


图8-Qwen-4B模型运行结果图

Qwen-4B微调后的模型输出效果如下：

graph TD A["细胞分裂专题"] --> B["分裂原因"] B --> C["相对表面积减小"] B --> D["遗传物质不均衡"] B --> E["DNA复制"] A --> F["分裂过程"] F --> G["观察染色体变化"] G --> H["前期：染色体出现"] G --> I["中期：染色体排列中央"] G --> J["后期：染色单体分开"] G --> K["末期：染色体变回染色质"] A --> L["基因与性状"] L --> M["性状由基因控制"] M --> N["多基因控制性状"] L --> O["染色体组概念"] O --> P["染色体组定义"] P --> Q["含本物种一整套遗传信息"] O --> R["染色体功能差异"] R --> S["不同染色体控制不同性状"] L --> T["21三体综合征案例"] T --> U["多一条染色体导致异常"]

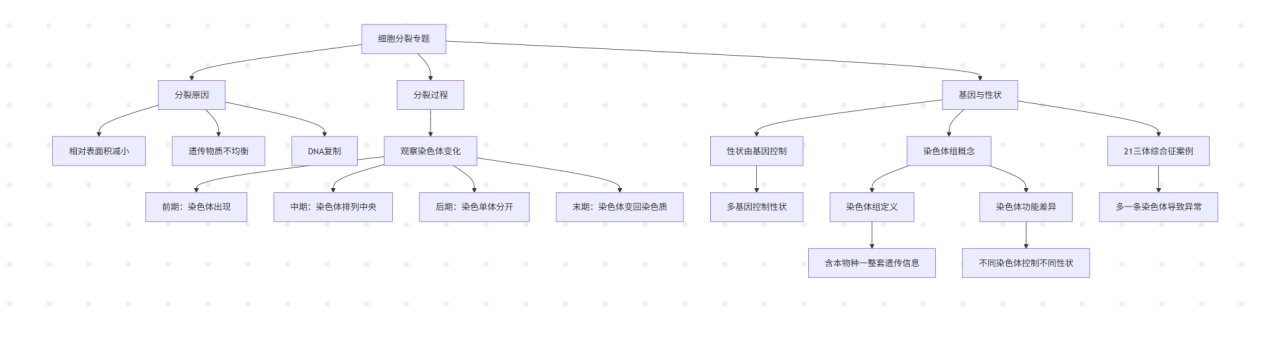


图9-Qwen-4B微调后模型运行结果图

Qwen-8B微调后模型输出效果如下：

graph TD O[细胞分裂专题] O-->A[细胞分裂的原因] A --> B[相对表面积减小] A --> C[蛋白质和酶需求增加] A --> D[遗传物质控制蛋白质合成] A --> E[细胞质不均衡] O-->F[细胞分裂的过程] F --> G[前期] F --> H[中期] F --> I[后期] F --> J[末期] O-->K[基因和性状] K --> L[染色体组] K --> M[基因控制性状]

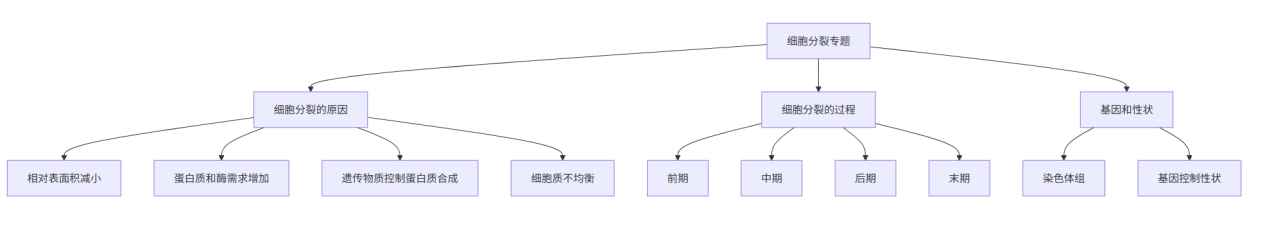


图10-Qwen-8B微调后模型运行结果图

可见，思维导图都是树形式的数据结构。我们认为树的深度能够体现模型对一个话题完整信息提取的能力，同时，树的分支数越符合原语料的分点数，越能体现模型的理解准确性。当有相同子树数的情况下，树越深，越体现出模型能够对一个话题信息进行较完备的提取。如上图各模型运行结果所示，Qwen-4B微调后的模型在深度上优于微调前的Qwen-4B模型，同时也更接近于参数量更大的deepseek-671B模型与Qwen-32B。在分支数上，微调前的Qwen-4B模型分支过多，体现了对语义把握的不透彻，而微调后的Qwen-4B模型在分支数上与语料本身契合，与参数量更大的deepseek-671B模型与Qwen-32B更为相似。此外，Qwen-4B微调后的模型与微调后的Qwen8B模型效果相似，但考虑到Qwen4B速度更快，最终部署采用Qwen-4B模型。

有关模型部署AI加速部分，对微调后的模型进行int4量化，上下文窗口平衡文本长度与速度后，选取5000，以应对大多数教学场景，部署后表现优异，启动仅需花费7s,token输出能达到10token/s。使用同样的prompt输出与量化部署前大体相同。

graph TD A["细胞分裂专题"] --> B["分裂原因"] B --> C["相对表面积减小"] B --> D["遗传物质不均衡"] A --> F["分裂过程"] F --> G["观察染色体变化"] G --> H["前期：染色体出现"] G --> I["中期：染色体排列中央"] G --> J["后期：染色单体分开"] G --> K["末期：染色体变回染色质"] A --> L["基因与性状"] L --> M["性状由基因控制"] M --> N["多基因控制性状"] L --> O["染色体组概念"] O --> P["染色体组定义"] P --> Q["含本物种一整套遗传信息"] O --> R["染色体功能差异"] R --> S["不同染色体控制不同性状"]

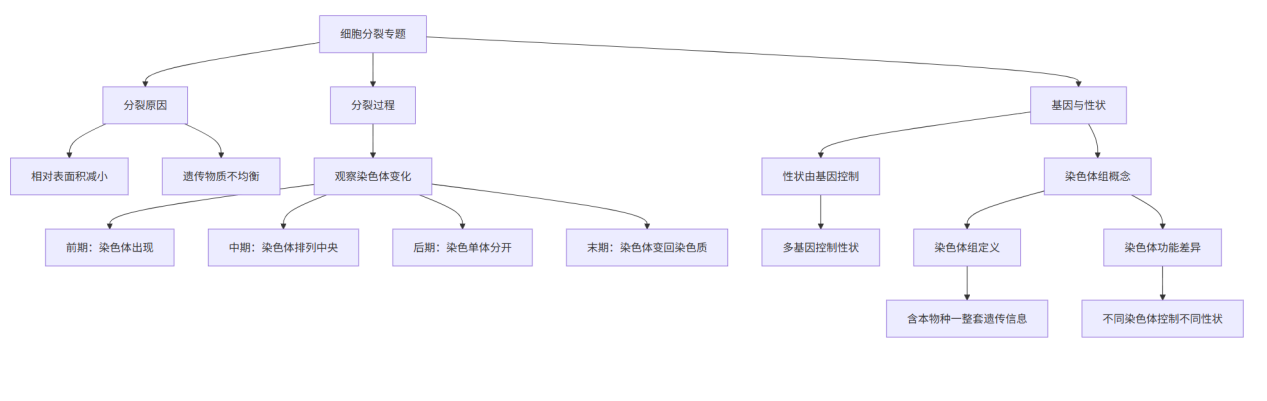


图11-微调后Qwen-4B模型部署运行结果图

# 第5章 成果

系统运行流程如下：教师在授课过程中，Airbox通过麦克风采集其语音输入将语音数据的Whisper模型进行语音识别，之后将转写得到的文本输入至微调的Qwen3-4B模型中，通过提取知识点、判断逻辑关系生成知识结构思维导图代码。生成的思维导图通过WebSocket协议实时推送给连接的学生端。学生在客户端可以查看导图、点击任意节点调用教材索引引擎，快速定位教材对应章节并高亮显示相关内容，同时学生还可通过客户端与教师进行实时互动与答疑。

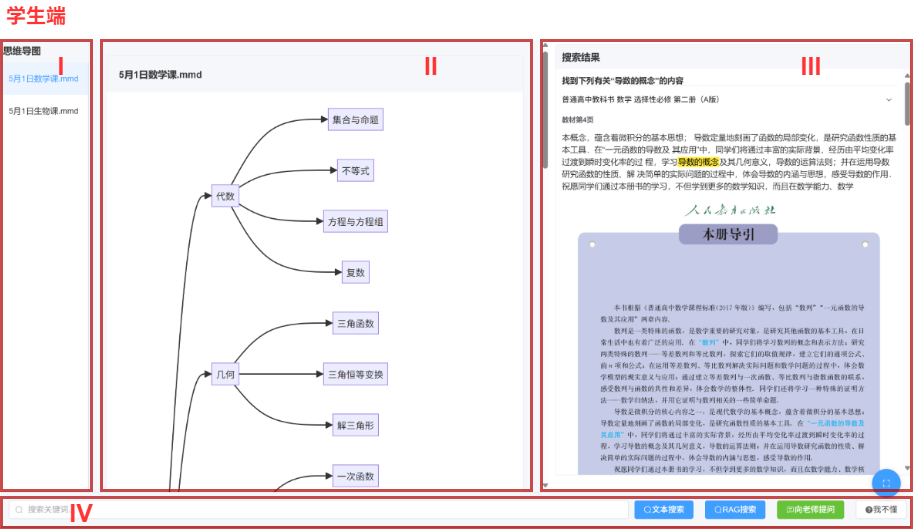


图12-客户端（学生端）界面示意图

学生端如上图所示，分为左中右下四部分，分别标号I、II、III、IV。

I区是思维导图列表。在这个区域，用户可以浏览不同的思维导图文件，并通过点击选择具体的文件来查看其内容。例如图中显示“5月1日数学课”“5月1日生物课”的思维导图等。

II区是当前选中的思维导图内容的详细展示区。在这里，用户可以详细查看历史记录中所选思维导图文件的内容，包括各个主题及其子主题的详细信息。此外，II区还支持思维导图实时查看，将实时生成的思维导图代码进行实时渲染，方便学生与老师在授课途中进行查看。图中展示的是“代数”部分的详细内容，包括集合与命题、不等式、方程与方程组、复数等思维导图节点，如对节点内容存在困惑与疑问，用户可以直接点击进行搜索，搜索内容将在III区进行进一步展示，进而帮助学生更深入地理解每个概念的具体内容和相互关系。

III区是搜索结果展示区。系统会显示与关键词相关的搜索结果。图中展示的是“导数的概念”的搜索结果，系统找到了相关的教材内容，包括导数的定义、性质和应用等，帮助用户快速定位和学习所需的数学知识。

IV区是互动区。用户可以在这个区域输入文本进行文本向量搜索或RAG搜索，系统会显示相关的搜索结果，如图13所示，左侧为基于课本知识点构建的图谱，右侧为AI针对学生问题的回答，学生在查看AI答案的同时，可以查阅AI给出该答案基于的具体的知识点，让AI解答不再是空中楼阁而是有理可依。同时，学生还可以点击“向老师提问”按钮向正在上课的老师实时进行提问。同时，如果学生跟不上老师的教学节奏，可以点击“我不懂”按钮提醒老师放慢节奏，重新解释一下学生不懂的内容。

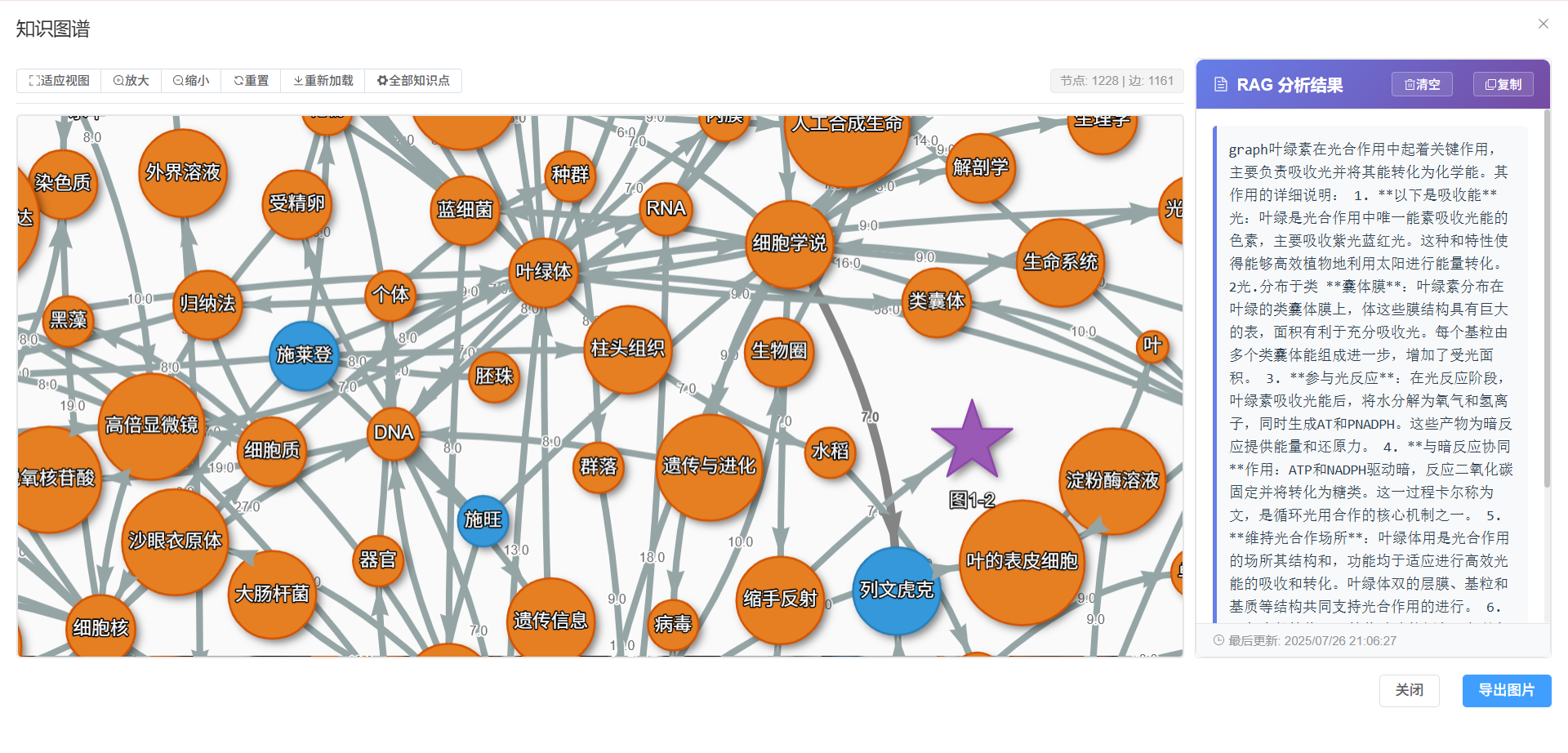


图13-客户端（学生端）RAG搜索界面示意图

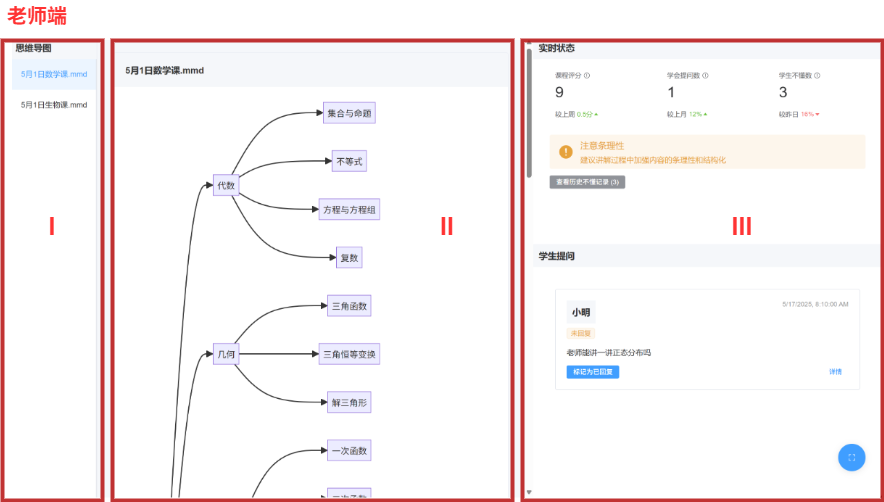


图14-客户端（老师端）界面示意图

老师端如上图所示，分为左中右四部分，分别标号I、II、III。

其中I、II区与学生端作用类似，不再赘述。III区是学生提问区域，它的作用是为老师提供一个可以查看学生问题与听课状态的区域。同时，教师也可以通过这个区域了解学生的学习需求和困难点，以便更好地调整教学策略和内容。图中显示了一个示例问题，学生“小明”询问“老师能讲一讲正态分布吗”，同时显示学生不懂数为3,示意刚才的部分有部分同学没有听懂，老师可以放缓授课节奏。

点击课堂评分，进入课堂评价页面，分为两层级评价，分别是内容评价和语言评价。

****

图15-课堂评价界面示意图

课堂内容评价界面如上，该界面简洁的展示了系统对于教师课堂内容不同方面的评分，教师可以参考这些数据调整授课节奏和授课策略，从而改善教学质量。

****

图16-语言习惯评价子界面示意图

上图为课堂评分界面的语言分析子界面。该界面会统计录音文本的口头蝉并对文本整体进行评分。该功能有助于帮助教师快速找到其语言习惯的不足之处从而改进授课质量。

# 第6章 创新点

**一、 基于课堂语音的实时知识图谱构建技术**

本系统依托Qwen3-4B微调大模型强大的语义理解能力，能够实时解析教师讲课内容中的关键知识点与逻辑关系，自动构建结构化的知识图谱。学生通过图谱可直观掌握知识框架，厘清复杂概念之间的联系，有效提升理解效率。同时，该图谱也为教师提供了教学过程的可视化分析工具，便于动态调整教学策略，实现因材施教。

1. **基于课本知识网络的智慧检索**

系统构建了与教材深度对齐的知识网络，并结合大模型RAG技术，打造了一套智能化的课本级检索机制。该机制不仅支持学生围绕具体知识点进行精准问答，更能从知识体系的角度提供完整、有逻辑的解答路径。与传统检索不同，这种基于知识网络的智能问答为学生提供“解题+溯源”的综合型学习支持，助力深度学习与迁移应用。

1. **端侧计算支持下的全流程数据本地化与隐私保护机制**

系统采用端侧计算架构，在无网络或网络受限的环境中亦可稳定运行。所有语音识别、图谱生成、数据分析等关键处理流程均在本地完成，不依赖云端服务器，极大地增强了数据安全性与隐私保护能力。这种全流程本地化部署不仅符合教育场景下对隐私合规的严格要求，也提升了系统的鲁棒性和可达性，为广泛部署提供了技术基础。

1. **开创教师备课多维评分新范式**

传统的教学评估往往基于经验丰富教师的听课评估、对比评估等手段进行，效果难以量化、缺乏稳定性且难以高频率进行。本项目设计的多维度评分体系可以给教师提供课堂质量相关评估作为参考更高效准确的改善课堂质量。评估总体由语言评分、内容评分等部分组成。语言评分使用统计、NLP等方法辅助教师修正口头蝉、冗余描述和语言风格问题；内容评分使用GNN识别课堂内容的导图，对课堂整体结构从内容、逻辑、层次等方面进行评估。相比传统方法，本系统的多维度评估可更频繁精确的完成，从而以帮助老师高效快速的改善教学质量。

1. **重构师生互动的智能教学新范式**

系统通过语音识别与思维导图实时生成技术，开创了全新的课堂交互模式。教师讲课内容被即时转化为结构化导图，学生可边听边看，理清知识脉络，增强吸收效率。学生可基于导图发起提问或请求教材链接内容，教师端实时接收并反馈，形成“讲-学-问-答”闭环。同时，“我不懂”按钮使教师能即时感知学生学习状态，调整节奏与难度，真正实现以学生为中心的个性化教学体验。同时，思维导图也可帮助教师梳理讲课条理性，提升教学质量。这种范式显著提升了课堂互动性与教学精准度，是智能教育领域的重要突破。

# 参考文献

1. Detmers, T., Pagnoni, A., Holtzman, A., & Zettlemoyer, L. (2023). QLoRA: Efficient Finetuning of Quantized LLMs. arXiv. <https://doi.org/10.48550/arXiv.2305.14314>
2. Hu, P., Lu, M., Wang, L., & Jiang, G. (2023). TPU-MLIR: A Compiler for TPU Using MLIR. arXiv. <https://doi.org/10.48550/arXiv.2210.15016>
3. Yang, An, Anfeng Li, Baosong Yang, and others. "Qwen3 Technical Report." arXiv, 14 May 2025. <https://doi.org/10.48550/arXiv.2505.09388>.
4. Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision. arXiv. <https://doi.org/10.48550/arXiv.2212.04356>
5. Evrekli, K., Ind, E., & Bahn, M. (2010). Scoring criteria of mind maps. Journal of Educational Technology & Society, 13(4), 123-135
6. FlowLearn: Evaluating Large Vision-Language Models on Flowchart Understanding. (2024). arXiv preprint arXiv:2407.05183
7. TEACHING AND ASSESSING MIND MAPS - Per Linguam. (n.d.). Per Linguam
8. 徐惠玲,刘城豪,潘苏东.数字化导向的职前教师评价素养培养研究[J].上海教育评估研究,2025,14(03):74-79.DOI:10.13794/j.cnki.shjee.2025.0038.
9. 陈丽.成果导向培训模式：基于中小学校长与教师培训实践的创新探索[J].中国教育学刊,2025,(08):85-90

.

附录

## 生成GNN训练数据使用的fewshot prompt：

1. 你是一个高中教学专家，精通使用Mermaid语法创建教学思维导图。
2. 你的任务是为以下指定的高中知识点生成一个Mermaid流程图代码，并根据详细的评分标准对其进行自我评估打分。
3. \*\*任务详情:\*\*
4. \* \*\*学科:\*\* {subject}
5. \* \*\*知识点:\*\* {knowledge\_point}
6. \* \*\*思维导图类型:\*\* {map\_type}
7. \*\*请严格遵守以下规则:\*\*
8. 1. 所有生成的Mermaid代码必须以 `graph TD` 开头，表示从上到下的布局。
9. 2. 生成的代码和打分必须\*\*完全符合\*\*下方评分指导中针对'{map\_type}'类型所描述的结构特点、内容深度和分数范围。
10. ---
11. \*\*评分指导 (满分: 100分)\*\*
12. 本思维导图的评分将从以下四个核心维度进行评估：
13. 1. \*\*结构逻辑性 (Structure & Logic) - 35分\*\*
14. \* \*\*详细型\*\*: 30-35分 (呈现网状或树状结构，多角度、多层次连接，逻辑紧密，知识覆盖广阔)
15. \* \*\*逐渐展开型\*\*: 25-30分 (树状结构清晰，有适当子分支，逻辑流畅，覆盖核心方面)
16. \* \*\*简略大纲型\*\*: 20-25分 (仅呈现基础主干分支，缺乏深入关联，覆盖最基础方面)
17. \* \*\*直线型\*\*: 15-20分 (节点单线联系，无法体现并行或层级，逻辑线性，仅覆盖单一维度)
18. 2. \*\*内容完整度 (Content & Completeness) - 35分\*\*
19. \* \*\*详细型\*\*: 30-35分 (包含定义、案例、数据、分析等多层次深度信息，概念准确，关键词精准)
20. \* \*\*逐渐展开型\*\*: 25-30分 (对核心概念有进一步解释和细化，深度有限，核心概念准确，关键词恰当)
21. \* \*\*简略大纲型\*\*: 20-25分 (节点多为关键词或短语，缺乏详细内容，内容准确性无法深入评估，关键词宽泛)
22. \* \*\*直线型\*\*: 15-20分 (内容概括，几乎没有深度，信息过少难以判断准确性，内容几乎是普通句子)
23. 3. \*\*层级清晰度 (Hierarchy & Clarity) - 20分\*\*
24. \* \*\*详细型\*\*: 18-20分 (拥有3个或以上清晰层级，主次分明，布局均衡，可读性高)
25. \* \*\*逐渐展开型\*\*: 15-18分 (拥有2-3个清晰层级，有效展示递进关系，布局基本均衡，可读性好)
26. \* \*\*简略大纲型\*\*: 12-15分 (仅有2个层级，结构扁平，布局简单，可读性尚可)
27. \* \*\*直线型\*\*: 10-12分 (伪层级，本质线性，无法体现层级，布局单调)
28. 4. \*\*代码规范性 (Code & Syntax) - 10分\*\*
29. \* \*\*详细型\*\*: 8-10分 (语法无误，可读性优秀，适当使用高级语法)
30. \* \*\*逐渐展开型\*\*: 7-8分 (语法无误，可读性良好)
31. \* \*\*简略大纲型\*\*: 6-7分 (语法无误，可读性一般)
32. \* \*\*直线型\*\*: 5-6分 (语法无误，可读性一般)
33. \*\*总分指导:\*\*
34. \* \*\*详细型\*\*: 90-100分
35. \* \*\*逐渐展开型\*\*: 80-89分
36. \* \*\*简略大纲型\*\*: 70-79分
37. \* \*\*直线型\*\*: 60-69分
38. ---
39. \*\*示例 (请仔细参考这些示例的风格、结构和评分):\*\*
40. \*\*示例 1: 详细型思维导图\*\*
41. \*\*学科:\*\* 生物
42. \*\*知识点:\*\* 光合作用
43. \*\*类型:\*\* 详细型
44. ```mermaid
45. graph TD
46. A["光合作用"] --> B["光反应阶段"];
47. A --> C["暗反应阶段"];
48. A --> D["总反应式"];
49. B --> B1["地点: 类囊体薄膜"];
50. B --> B2["条件: 光、色素、酶"];
51. B --> B3["物质变化"];
52. B --> B4["能量变化"];
53. B3 --> B3a["水的光解: 2H₂O → 4[H] + O₂"];
54. B3 --> B3b["ATP的生成: ADP + Pi + 能量 → ATP"];
55. B4 --> B4a["光能 → 活跃的化学能 (储存在ATP和[H]中)"];
56. C --> C1["地点: 叶绿体基质"];
57. C --> C2["条件: [H]、ATP、多种酶"];
58. C --> C3["物质变化"];
59. C --> C4["能量变化"];
60. C3 --> C3a["CO₂的固定: CO₂ + C₅ → 2C₃"];
61. C3 --> C3b["C₃的还原: 2C₃ + [H] + ATP → (CH₂O) + C₅"];
62. C4 --> C4a["活跃的化学能 → 稳定的化学能 (储存在糖类中)"];
63. D["6CO₂ + 12H₂O --(光/叶绿体)--> C₆H₁₂O₆ + 6O₂ + 6H₂O"];
64. code
65. JSON
66. {{
67. "Structure\_Logic": 34,
68. "Content\_Completeness": 35,
69. "Hierarchy\_Clarity": 19,
70. "Code\_Syntax": 9,
71. "Total\_Score": 97
72. }}
73. 示例 2: 逐渐展开型思维导图
74. 学科: 生物
75. 知识点: 光合作用
76. 类型: 逐渐展开型
77. code
78. Mermaid
79. graph TD
80. A["光合作用"] --> B["光反应阶段"];
81. A --> C["暗反应阶段"];
82. B --> B1["地点: 类囊体薄膜"];
83. B --> B2["产物: [H], O₂, ATP"];
84. B2 --> B2\_Detail["为暗反应提供[H]和ATP"];
85. C --> C1["地点: 叶绿体基质"];
86. C --> C2["过程: CO₂的固定与C₃的还原"];
87. C2 --> C2\_Detail["消耗[H]和ATP，合成糖类"];
88. code
89. JSON
90. {{
91. "Structure\_Logic": 28,
92. "Content\_Completeness": 29,
93. "Hierarchy\_Clarity": 17,
94. "Code\_Syntax": 8,
95. "Total\_Score": 82
96. }}
97. 示例 3: 简略大纲型思维导图
98. 学科: 生物
99. 知识点: 光合作用
100. 类型: 简略大纲型
101. code
102. Mermaid
103. graph TD
104. A["光合作用"] --> B["光反应"];
105. A --> C["暗反应"];
106. A --> D["影响因素"];
107. code
108. JSON
109. {{
110. "Structure\_Logic": 22,
111. "Content\_Completeness": 21,
112. "Hierarchy\_Clarity": 14,
113. "Code\_Syntax": 7,
114. "Total\_Score": 74
115. }}
116. 示例 4: 直线型思维导图
117. 学科: 生物
118. 知识点: 光合作用
119. 类型: 直线型
120. code
121. Mermaid
122. graph TD
123. A["吸收光能"] --> B["水的光解"] --> C["ATP合成"] --> D["CO₂固定"] --> E["C₃还原合成糖类"];
124. code
125. JSON
126. {{
127. "Structure\_Logic": 16,
128. "Content\_Completeness": 18,
129. "Hierarchy\_Clarity": 10,
130. "Code\_Syntax": 6,
131. "Total\_Score": 60
132. }}
133. 期望的输出格式:
134. 请严格按照 "Mermaid代码块 + JSON评分块" 的格式输出，不要有任何额外文字。
135. 首先，输出Mermaid代码块，以 mermaid 开始，以 结束。
136. 在代码块之后，提供评分细节，采用以下JSON格式:
137. code
138. JSON
139. {{
140. "Structure\_Logic": [score],
141. "Content\_Completeness": [score],
142. "Hierarchy\_Clarity": [score],
143. "Code\_Syntax": [score],
144. "Total\_Score": [total\_score]
145. }}